User Charges for Health Care: A Review of the Underlying Theory and Assumptions

Germano Mwabu
User Charges for Health Care: A Review of the Underlying Theory and Assumptions

Germano Mwabu
UNU/WIDER

March 1997

This study has been prepared within the UNU/WIDER project on New Models of Provision and Financing of Public Goods in Developing Countries, which is co-directed by Germano Mwabu, Senior Research Fellow, and Reino Hjerpe, Director General of the Finnish Government Institute for Economic Research.

UNU/WIDER gratefully acknowledges the financial contribution to the project by the Government of Sweden (Swedish International Development Cooperation Agency - Sida). The author would like to thank Sudhir Anand, Robert Evenson, Nguyuru Lipumba, E. Wayne Nafziger and Sanjay Reddy for their helpful comments.
### CONTENTS

**ABSTRACT**

**I INTRODUCTION**

1.1 Health and economic growth 1  
1.2 Health care financing concerns 2  
1.3 Structure of the paper 4

**II THE THEORETICAL BASIS FOR USER CHARGES**

2.1 Efficiency results of the price system 5  
2.2 Market-based reforms and user fees for health care 7  
2.3 User fees as efficiency signals and the distributional concerns 7  
2.4 The policy value of assumptions 9

**III CAVEATS ON USER CHARGES**

3.1 Regularity of production technologies and of consumer preferences 10  
3.2 Existence, completeness and imperfectness of markets 12  
3.3 Information asymmetry and uncertainty 15  
3.4 Institutions, transactions costs and scale effects 20  
3.5 Coordination and regulation of actors in the health sector 31

**IV MEDICAL CARE DEMAND AND THE IMPLEMENTATION OF USER CHARGES IN DEVELOPING COUNTRIES**

4.1 Issues and assumptions 41  
4.2 Price elasticity of the demand for medical care 43  
4.3 Medical care quality 46  
4.4 Fee implementation procedures and health service management 46

**V THE HEALTH CARE FINANCING GAME**

5.1 Preliminaries and the nature of the game 49  
5.2 The solution concepts 51  
5.3 An illustrative application of rationalizability 53  
5.4 The coalition of players and the special attributes of civil-society 54

**VI SUMMARY AND CONCLUSIONS**

6.1 Summary 57  
6.2 Conclusions and policy implications 59

**REFERENCES** 63
ABSTRACT

The paper reviews the theoretical basis for the application of user fees in the public health sector in low-income countries with particular reference to the special characteristics of medical care as a commodity. The general equilibrium efficiency result of the market mechanism is shown to be the theoretical justification for the financing of health services via a system of user charges. If markets for all goods and services exist, and are perfect in a very strict sense, the welfare outcome of the price mechanism cannot be improved upon by any other resource allocation device. Furthermore, the decentralized and impersonal nature of this mechanism renders it more convenient to use in the allocation of commodities, health care included, than its alternatives such as a system of centrally administered prices or a system of administrative controls and directives. However, since many of the assumptions of the price system are rarely met in actual situations, especially in the health sector, it should be applied with caution.

In particular, problems of information asymmetry and consumption externalities in health care markets necessitate a simultaneous use of fees with government interventions in order for fees to achieve their often intended aim of efficiency and equity improvement in health care provision. The most important intervention of the government here is the enactment and enforcement of institutions that reduce costs of transacting in health care markets and that in addition facilitate the emergence of new markets such as the markets for medical insurance.

A striking finding of the paper is that health services in low-income countries are best financed primarily by revenue from general taxation, supplemented by a system of moderate user fees. Since medical insurance markets are generally non-existent in low-income areas, it is argued that financing health services primarily through user fees in such areas would be inefficient and inequitable. However, to mitigate the moral hazard problem as well as the problem of the commons, both of which characterize publicly financed health care, imposition of modest user fees is required. The importance of fees in this proposal increases with economic growth and with evolution of institutions that facilitate market transactions.

Strategic interaction among economic agents is shown to affect the structure and implementation of user fees. A game-theoretic analysis of the general problem of health care financing shows that this problem is best
tackled by harnessing the efforts of households, private health care providers, the government and civil society. These entities form what might be called a winning coalition in health care financing game of society. It is argued that the government is better placed to provide an institutional framework for coordinating the efforts of the various players to the desired end.
I INTRODUCTION

1.1 Health and economic growth

Over the past ten years, research on economic growth has demonstrated that human capital is a powerful force in the development process (see e.g., Romer 1986; Lucas 1988; Becker 1990). In consequence, a sustained increase in this form of capital is crucial for poverty reduction in low-income countries and for an ever rising standard of living. Education and health are two commonly used proxies for human capital – an unobservable magnitude or force that is part and parcel of human beings (Schultz 1960, 1961, 1963).

However, there is an issue as to whether better health is the consequence or the cause of economic growth. Pritchett and Summers (1996) argue that wealthier nations are healthier nations. They attribute better health to economic growth, and show with aggregate data from developing countries that this association is not incidental, but causal and structural, using the latter term to mean that for a given level of income, the structure of an economy does influence the observed health status, as measured for example, by life expectancy or infant mortality rate. Benefo and Schultz (1994, 1996) obtain the same result with micro data from Ghana, also using instrumental variable estimation approach. This result, however, does not imply that the reverse causation is absent from the estimated health-income relationship; all it says is that simultaneity bias has been accounted for in the estimation of the effect of income on health. If instead, the focus were on the measurement of the effect of health on income, i.e., on reverse regression, the same estimation strategy would apply. Thus, no damage has been done to wage or growth regressions with human or health capital as the conditioning variable. The motivation of this study is to explore effective ways of financing health investment, both for its own sake as a consumption good, and as a tool for improving labour productivity and economic growth generally.

A strong positive correlation between education and health in survey data has recently been interpreted to imply that these two variables are aspects of the same thing, i.e., investment in human capital, rather than that one variable causes the other (see Fuchs 1996, pp. 5-6). It follows therefore that an increase in health investment, other things being equal, would increase the available stock of human capital – a force that is conducive to
economic growth and hence to a general improvement in social welfare. Two factors are particularly important in increasing the available stock of health capital: availability of technology to produce additional health investment and the financial resources to pay for that investment. Assuming that health augmenting technology exists, the key issue to be settled is the choice of the modality for financing the additional health investment. Health financing has been a most controversial and difficult issue to tackle in virtually all developing countries over the past decade (see e.g., Reddy and Vandemoortele 1996). Technologies and facilities to improve health status exist in these countries and the start-up capital needed to finance their acquisition could be made available via development aid and/or external debt. The problem has been how to meet the recurrent cost of using these technologies to make health maintenance services broadly available to the population on a sustained basis. This financing problem arises from six different aspects of health care.

1.2 Health care financing concerns

The first aspect has to do with the special nature of medical care as a consumer good. Health care, in curative or preventive form, is widely regarded as a merit good – a commodity that ought to be available for use by everyone irrespective of ability to pay. Ideally, the 'merit good' attribute constrains the financing method in usage to be non-exclusionary, in the sense of not preventing any population group from consuming health services of acceptable quality. Needless to say the non-exclusionary constraint is quite hard for most low-income countries to fulfil.

The second dimension of the financing problem arises due to with social externalities in the consumption of health care, i.e., to social effects of private health care consumption, such as the public protection from infectious diseases due to health maintenance activities of individuals. These effects call for the design of a financing mechanism that gives private agents incentives to consume the socially optimal quantity of preventive and curative health services. Such a mechanism is difficult to design because it requires information on individual preferences as well as on the ability to measure benefits derivable from health care use. Thirdly, some forms of health care, such as immunizations, are a public good in the sense that the benefit from these services – protection from infectious diseases – is non-excludable and thus a case can be made for their public financing. Note however that since the services themselves are excludable
in consumption, there may be efficiency grounds for their private financing.

The fourth difficulty in the design of a health care financing method relates to the fact that curative health care is unpredictable so that it might be needed when it cannot be afforded. For this reason, medical insurance – an institutional arrangement for paying for medical care in advance ought to be an integral component of a medical care delivery system. Development of insurance schemes in low-income countries however is not an easy task because institutions for risk bearing are limited. Moreover, even when such institutions exist, the required premiums might not be affordable by the majority of the population. This is especially likely to be the case in rural areas where the bulk of household income may not be in form of cash or in a form that can easily be converted into cash.

The fifth problem in health care financing in low-income countries is related to the limited nature of the tax base in these countries so that additional revenue for the health sector from the government can, in the short term, mainly be obtained by restructuring or reallocating the budget rather than via extra taxation. It is this fact that often makes user charges appear as if they are the only feasible alternative source of additional finance for health sectors in low-income countries. As is shown in a later section of the paper, presumptive taxes and changes in property rights, can considerably expand the range of potential sources for health finance in many developing countries. The sixth problem is related to information asymmetries in health care markets and in the related market for health insurance. This information asymmetry has certain implications for the pricing of health care services and for the coverage of health risks. The first basic point in this regard is that since the market information is imperfect, health service prices might not reflect the quality of the services being offered by private agents. A health care financing method based primarily on a system of user fees must therefore provide for an active role for government intervention in private health care markets to ensure that patients are not charged excessive fees or given services of insufficient quality. The second and related point is that due to information and/or institutional problems, markets for certain types of health services might not exist, in which case, health service provision and financing should be the responsibilities of public authorities.
Thus issues of efficiency, equity, information asymmetry, publicness, social externalities, provider regulation and coordination, as well as matters of property rights should be carefully considered when designing a health care financing system. Previous discussion on user charges for health care has mainly been limited to issues of efficiency and equity and only within the health sector. This paper is an attempt to rectify this unfortunate situation. It argues that health care financing policy should simultaneously consider the budgetary and behavioural effects of user fees in the context of health care and related markets. Further, the behavioural or welfare effects of fees for health care should be evaluated considering also the effects of fiscal measures, medical insurance, and the institutional structure of society. These matters are important aspects of public policy on health care financing and need thorough study.

1.3 Structure of the paper

This paper has five sections following this introduction. The second section presents the price system as the theoretical basis for user fees in the health sector in a market economy. The third section provides a detailed discussion of the assumptions that underlie the efficiency results of the market mechanism, relating them to institutional and structural features of health care sectors in developing countries. The section reveals the central role of the government in an ideal provision and financing of health services in a market economy. This role arises from the power of the government to create and enforce institutions that facilitate and coordinate interaction of health care providers and consumers in imperfect health care markets. The fourth section discusses the prominent role ascribed to the magnitude of price elasticity of the demand for medical care in implementing user fees in developing countries; it is shown that the emphasis placed on price elasticities in the setting and implementation of fees was unwarranted. Issues of incentive structure in public health sector, health service quality and management were at least as important. In section five, the consequence for health care financing of strategic interaction among economic agents is examined. The problem of health care financing is viewed as arising from the need to raise funds (through fees, taxes or other mechanisms) to pay for treating illnesses which are assumed to be randomly and repeatedly inflicted to the population by an 'invisible hand' of nature. In consequence, health care financing is envisaged as a repeated game of strategy played against nature by a coalition of households, health care providers, the central government and
civil society. It is shown that civil society is a crucial member of this coalition. Section six summarizes and concludes the paper.

II THEORETICAL BASIS FOR USER CHARGES

2.1 Efficiency results of the price system

Two basic results of neoclassical analysis underlie the policy of using market prices (rather than administrative directives) to allocate goods and services in an economy, namely: the first and the second fundamental theorems of welfare economics. The first fundamental theorem of welfare economics states that under certain conditions (to be spelt out shortly) the market allocation of goods and services is Pareto efficient at competitive equilibrium. That is, resource allocation is such that it cannot be changed to make one person better off without making someone else worse off, a situation which implies that from an individualistic standpoint, the best possible use has been made of the available resources. Furthermore, since a whole range of competitive market equilibria is possible, each such equilibrium is Pareto-efficient. Accordingly, the second fundamental theorem of welfare economics states that any Pareto-efficient equilibrium can be obtained through the use of the price mechanism – that is, public authorities can manipulate the market system (for example by lump-sum subsidies) to achieve efficiency in resource allocation for any desired distribution of wealth.

The foregoing efficiency results are based on five key assumptions. First, the economy is assumed to be market-based and competitive: consumption and production decisions are made by numerous, self-regarding consumers and producers, none of which has the power to influence market prices. The second and related assumption is that resource endowments are privately owned so that consumption and production activities are decentralized to numerous households and firms. Third, all market participants have perfect foresight about future consumption and production plans and are in addition perfectly informed about characteristics of the goods and services being traded in all spot and futures markets, e.g., about service quality and location. Fourth, there exists a complete set of markets for all goods and services, including markets for risk coverage. Fifth, production technologies and consumer
preferences are regular, e.g., there are no indivisibilities or externalities in production and/or consumption.

The above assumptions constitute an idealized market-based economy, with the property that there exists an equilibrium set of prices at which resource allocation is Pareto-efficient (Arrow and Debreu 1954). The role of market prices in this highly decentralized system is to coordinate activities of economic agents so that at equilibrium, commodity supplies and demands are simultaneously equal in every market. The main contribution of Arrow and Debreu (1954) was to show that under the assumptions just stated, such an equilibrium exists and is constrained Pareto efficient. It is also known from Lange-Lerner-Taylor Theorem (see Stiglitz 1995) that even in an economy characterized by public ownership of factors of production, a Pareto-efficient allocation of resources can be achieved through the use of prices.

The policy conclusion of Lange-Lerner-Taylor Theorem coincides with, and strengthens that of the second fundamental theorem of welfare economics: irrespective of whether property rights are publicly or privately owned, the price system can be used to obtain any Pareto-efficient outcome in society. As instruments for resource allocation, prices carry two pieces of information. First, they convey to market participants information about resource scarcities, i.e., information concerning the structure of returns or benefits derivable from the available resources. Market prices are a form of an incentive structure available freely to households and to firms for efficient use of resources because they signal the true benefit or cost of using each resource. Second, prices that households are willing to pay for commodities convey information to firms about the kinds of commodities consumers want. Further, the prices at which firms are willing to sell commodities transmit to households the costs of producing the commodities the households desire to consume. The great merit of competitive prices is that they reflect marginal benefits of consuming commodities as well as the marginal costs of production so that their coordination of production and consumption decisions of firms and households is perfect: given these prices, firms efficiently produce exactly the commodities consumers want and in the right quantities. The informational and coordinating role of the price system is neither restricted to an economy characterized by private ownership of property rights, as in the Arrow-Debreu model, nor to an economy in which productive resources are publicly owned, as in Lange-Lerner-Taylor model. The
coordinating role of prices is also present and operative in a mixed economy – that is, an economy characterized by public as well as by private ownership of factors of production. Moreover, to the extent that the standard neoclassical assumptions hold, the allocative outcome of the price system in a mixed economy is constrained Pareto-efficient.

2.2 Market-based reforms and user fees for health care

It is now a short step to the connection between the standard neoclassical theory of an economy, and the structural adjustment policy of using user charges to finance health care in the public sectors adopted by many developing countries over the past decade or so (see for example Reddy and Vandemoortele 1996). The connection rests on a simple observation that the market for health care is one of the many markets constituting an economy so that price reforms aimed at efficiency improvement should be applied in all markets. Accordingly, the policy of market liberalization (e.g., price decontrols and removal of price subsidies) embraced by many developing countries in mid-1980s and early 1990s was naturally extended to health care markets in the public sector, where health services were provided practically free of charges. Health services, like other goods and services, require scarce resources to produce. A mechanism for allocating these services, such as a system of user charges (a set of fees differentiated by type of service), which signals the scarcity of health care resources is conducive to efficiency both in service provision and consumption. It is thus more likely to yield greater health gains than a financing method without such an in-built efficiency incentive. Throughout this paper, we use the term 'user charges' or 'user fees' in the sense used by Reddy and Vandemoortele (1996) as 'any form of contribution to costs [of service provision] by users.' Needless to say, the contribution made by users might be equal to, below, or above the cost of service provision. For a distinction between cost recovery, cost sharing, user fees and community financing see (Reddy and Vandemoortele 1996).

2.3 User fees as efficiency signals and the distributional concerns

As to efficiency, a system of charges for health care carries information about the cost of using and providing health care. Regarding the information to users, a system of user fees signals to households what they must pay to obtain health services, and so it gives them an incentive to use the services well. The incentive lies in the fact that households face a hard
budget constraint so that spending part of this budget on unnecessary health care deprives them of benefits derivable from the consumption of other goods and services. It is because of the rational response by households to a hard budget constraint that user charges were seen in the earlier literature (see e.g., Griffin 1988) as reducing 'frivolous' demand for medical care – an efficiency enhancing effect. On the supply side, the cost information embedded in a system of user charges signals to providers the scarcity of the resources needed to produce health services. Under the assumptions already noted, such information is important in determining the resource bundle required to provide health services in the least costly way. In a manner analogous to that noted for households, both the hard budget constraint in the case of private providers, and the pressure of the central government to reduce the budget deficit in the case of public providers would induce cost conscious behaviour in service provision.

Consistent with the Second Fundamental Theorem of welfare economics, any distributional issues in the use of health services can be resolved using lump-sum income transfers. In view of this, and provided of course that the assumptions of neoclassical economics are reasonable approximations of real world markets, a system of user charges is an efficient mechanism for financing health services. Notice that in this case, private health sector is indistinguishable from the public sector: in both sectors, resources are efficiently allocated by the market mechanism. Public financing of health services in either sector would do no better than private financing in improving health status and hence, other things being equal, social welfare. Further, to the extent that public financing would involve imposition of taxes that would distort the structure of performance incentives in the economy, it would be inefficient.

Note that under idealized market conditions, a mixture of private and public methods of health care financing (the public funds are raised through lump-sum taxes) has the same welfare outcome as the market mechanism when used alone. That is, in this case government intervention in the financing of health care is unnecessary. However, when the actual market environment differs from the ideal one, a whole range of optimal mixtures of private and public mechanisms of financing health services presents itself. Moreover, it becomes necessary to distinguish between the private sector (where broadly speaking, service provision is by private agents or civil-society) and the public sector (where service provision is by the government). The optimal mixture chosen in that case depends on the
extent to which the actual and ideal market environments differ and on objectives of government. As to the policy effect of objectives, a government that emphasizes equity in its health care delivery system might favour a mixed financing strategy in which public financing is dominant. However, as argued previously, user charges do encourage rational use of health services. Thus even a government intent to use revenue from general taxation to finance health services in the public sector for equity reasons would have efficiency reasons for using a system of user charges as well. There is thus a strong theoretical basis in favour of user charges for health care in the public sector in both developed and developing countries.

2.4 The policy value of assumptions

The main point of the analysis so far has been to restate and put upfront the general equilibrium theory of markets on which the policy of private financing of health services through user charges is based. Since the theory is an idealized description of markets, its policy relevance is in providing a benchmark for choosing between alternative methods of financing health services in market economies. The policy conclusions drawn from this theory are not intended to be implemented as if the assumptions on which the theory rests hold in the real world. Instead, the policy value of any such conclusions lies in suggesting what can be done to make the performance of actual market economies or a specific sector of such economies, for example the health sector, approximate the ideal level of efficiency revealed by price theory and the assumptions on which it rests. It is important to stress that theory is only a guide to policy-making and policy implementation -- not a blueprint for these two activities both of which are highly intensive in personal judgement and in knowledge of local contexts in which they occur. Much attention is devoted to these issues in the ensuing section.

III CAVEATS ON USER CHARGES

There are persuasive theoretical arguments against exclusive or primary reliance on a system of user fees as a mechanism for financing and allocating health services, particularly in low-income countries. All the arguments revolve around a phenomenon known as the 'market failure' – the inability of health care markets, if left alone, to yield a Pareto-efficient
outcome in society. It should now be noted that an individualistic Pareto-efficient outcome of the price system can be very undesirable from the standpoint of society. However, the second fundamental theorem of welfare economics implies that such an undesirable situation can be corrected by a simple fiscal measure – a lump-sum redistribution of wealth among individuals without affecting the functioning of the price system (or in the particular case of health care markets without affecting the allocative efficiency of a system of user charges). The market failure argument states that even with a lump-sum redistribution of income, health care markets or markets in general would not be efficient because the postulates of price theory, on which the efficiency results are based, are an incorrect characterization of the real world economies (see Stiglitz 1995, especially chapters 3 and 4). In consequence, the fundamental theorems of welfare economics do not hold and thus government interventions in the provision and financing of health services are more warranted than is suggested by the standard market failure arguments.

From this point on, attention will be turned to real world aspects of health markets that are misrepresented or missed by basic assumptions of standard price theory. Recall this is the theory, as argued in section 2, that provides the rationale for a variety of systems of user charges implemented in many developing countries over the past decade as part of market-based economic reforms. Needless to say, implementation of this component of reforms is still in progress in many developing and transition economies. There is urgency therefore in articulating theoretical reasons for market failure in the health sector for the purpose of informing and reorienting user fee policies in these economies.

Several assumptions of price theory lead to wrong analytical conclusions concerning efficiency outcomes of markets, particularly the markets for health services, and hence to inappropriate policy recommendations. We focus attention on regularity, existence, information, institutional and coordination assumptions and examine health services financing implications of each in turn.

3.1 Regularity of production technologies and of consumer preferences

The regularity assumption concerns a broad spectrum of restrictions on production technologies and consumer preferences. For example, the
convexity assumption rules out the phenomena of indivisibilities and externalities in provision and utilization of health services. With regard to provision, particular types of health services such as hospital services require a lumpy investment to produce – an investment which individual providers would normally be unable to finance. Without government interventions, such services would be undersupplied by markets because providers would be unable to raise the required investment capital. A low-income community in a moderately-sized geographic region such as a district for instance, might go without hospital services (including facilities for training primary health care givers) because of indivisibilities in the provision of hospital care.

On the demand side, health services are characterized by extensive externalities in consumption (de Ferranti 1985). Further, certain outcomes of consuming particular kinds of health services, for example protection from infectious diseases via immunization, or a cure from such diseases after receiving medical treatment, are public goods whose benefits cannot be excluded from persons unwilling to pay for preventive or curative health services. If the financing of immunization services for example were to be left to the price system, they would be undersupplied and thus the efficiency theorems of welfare economics would not hold. That is, the market would fail to provide them in optimal quantities because of the free rider problem in their financing. Health services for which some individuals may have incorrect preferences (preferences inconsistent with social benefits of consuming such services) may also be supplied in suboptimal quantities. Examples of these include a variety of preventive health services and particular forms of curative care such as treatment for chronic illness. Individuals may be unwilling to pay for these services not because of a strategic consideration, as in the case of the free rider problem, but because they do not value benefits derivable from consuming them. Reliance on user charges to finance health services characterized by non-rivalrous consumption or for which personal preferences are defective would be inefficient. This does not imply that government financing of such services would be efficient. What is implied is that because of the great practical difficulties involved in setting appropriate user fees for such services efficiency losses may be smaller when publicly financed. In view of these considerations, health services that fall under the category of public or merit goods are better provided and financed by the government.
3.2 Existence, completeness and imperfectness of markets

The assumption that markets exist and are complete and perfect for all types of health services leads to a prediction of efficiency outcomes of markets that are not obtainable in real world situations. The obvious point to make in this connection is that markets do not exist for all kinds of medical and health services. Markets for immunization services or for dental care, for instance, do not exist in many rural areas of developing countries. Also absent in these areas are markets for health insurance – a crucial market in situations where health care is financed through user charges (Shaw and Griffin 1996). Further, even when markets for health care or insurance exist, they are incomplete in most cases, i.e., some attributes of the services provided – attributes which households value – such as service quality and continuity cannot be obtained via market transactions. Because of certain informational problems (see below) households or consumers may not be willing to pay for these attributes even when the market can provide them. Thus public financing of medical care may be preferred over a system of user charges because of incompleteness of markets; their public provision may also be warranted on similar grounds.

Government intervention in health care markets (broadly defined to include markets for health insurance) could further be justified on grounds of market imperfections, due for example to monopoly power. Economies of scale (characterized by a fall in average cost as the size of the unit of service provision increases) or economies of scope (associated with a fall in average cost as the range of services being provided in a given health unit increases) may lead to monopolistic health care markets. Exclusive reliance on user charges to finance health services in such markets would be inefficient because the services would be undersupplied; the user fees themselves would be excessive for they would exceed those that would prevail in a competitive perfect. Government intervention in form of subsidies or regulatory controls would in this instance mitigate the welfare losses arising from market imperfections. It is obviously unreasonable to treat a monopolistic health care market as if it were perfect in designing a health care financing strategy. An example of such an unreasonable policy, occasionally encountered in developing countries, is privatization of a national referral hospital (or a segment of it) and letting it free to set user charges for the services it provides. Other things being equal, the consequences of such a policy, which transforms a public monopoly (a
referral hospital) to a private one are higher service prices, a lower level of service provision and a drop in service demand.

An additional reason for government intervention in the financing of large-scale health care facilities, e.g., referral hospitals, is that private providers might fail to raise the capital needed to construct them and their benefits to society would thus be lost. It is important at this point to distinguish between government funding for construction of health facilities characterized by scale economies and government funding of service provision by such facilities. The appeal for public funding of costs of constructing large scale health facilities presumes imperfections in capital markets, whereas the appeal for government funding of their recurrent costs presupposes imperfections in credit markets (e.g., as evidenced by inability of creditworthy households to borrow to finance consumption, including health care). It should be noted here that the functioning of health services markets is necessarily being examined in relation to other markets, particularly the credit and capital markets. Also to be observed, is the fact that the issue of financing hospital services through user charges presumes the existence of hospitals, for without their existence, the issue would not arise. This is why the linkage between the market for hospital services and the market for financial capital is important for an understanding of the role of the government in the supply side of health services market, either as a direct provider of the services that are subject to scale economies or as a facilitator in the funding of private provision of such services. We neglect all the supply-side issues of this market and focus attention on its demand-side and the role of the state there. To start with, there is need to observe that even for a service which is subject to scale economies, user charges can be quite high in the initial stages of its provision, i.e., before the optimum level is reached. In fact, the charges might be so high that without subsidies from the government, people would not afford to pay the charges and service provision would be stopped. The well-known infant industry argument (IIA) applies here in full force.

Notice however that once the optimum service provision is reached, and the IIA ceases to be the basis for government intervention, policy makers would take a step farther back and use the earlier market imperfection reasoning to justify continued intervention. In both situations, the intervention can be via subsidies or regulation. In the case of a monopolist,
the regulatory intervention would control user charges, whereas in the case of an infant provider, it would be protection against competition.

There is yet another reason for justifying protection in the health sector. Contrary to the Arrow-Debreu model of ideal markets, Krueger (1974) in her pathbreaking paper has shown that restriction of competition might be efficiency-improving in a sector characterized by a competitive rent-seeking behaviour. The idea here can be illustrated with reference to the market for drugs in the private sector and the market for curative health services in the public sector. The level of user charges in the market for curative services depends crucially on prices at which the government can buy drugs from private sector. To obtain Krueger’s result for the health sector – i.e., the result that restricting competition might be helpful – assume that the government procures drugs by issuing contracts to private drug suppliers. Since the contracts issued by the government (which may include a license to import drugs) are scarce, there are rents to be earned by suppliers who get the contracts. It follows therefore that private drug suppliers will compete in a variety of ways (including bribery) for rents that can be earned by winning government contracts. The intensity and form of competition would depend on magnitudes of the expected rents. By restricting such competition, for example via establishment of a monopolist to supply drugs to the public health facilities, the government can reduce the efficiency loss due to rent-seeking behaviour. Needless to say, the resources saved by this measure can be used to produce other goods and services, which in addition to health care are important in improving household welfare. In fact, government central drug stores in many developing countries (which have the status of a monopolist) help mitigate, albeit unintentionally, the rent-seeking behaviour in the procurement of drugs. If the monopolist status also facilitates realization of scale economies in the procurement of drugs, it makes it possible to reduce user charges in public health facilities or to levy moderate fees in facilities where they do not exist. The lowering of fees enables the government to extend health services to a larger population without increasing the health budget. An important policy lesson to be drawn from this sub-section is that government efforts aimed at correcting distortions in health services markets are more likely to be effective if they take into account imperfections in related markets.
3.3 Information asymmetry and uncertainty

A problematic aspect of price theory is that all market participants are assumed to be equally- and well-informed about the commodity or service being traded. This assumption is too strong for health care markets. The issues involved here have been well-known for a long time in the health economics literature (Arrow 1963) and in general economics literature (Akerlof 1970). The issues have recently been articulated in great detail by Stiglitz (1995). The informational assumption of price theory in the context of health care market states that both households and health care providers are well-informed about the quality of health services. That is, the patient has the same information about medical care as the physician. Once this assumption is disregarded, the phenomena of information asymmetry and uncertainty in health care markets become important in analysing efficiency implications of user charges. To stress this point, the doctor knows much more about the quality of his services and his skill as a physician than the patient knows about these matters (see North 1990 for other examples of information asymmetry).

The behavioural implication of this informational asymmetry is that in consuming health services, a patient enters into an agency relationship with the provider. Assuming that the user fee faced by both the patient and the provider reflects the marginal cost of service provision, the implicit agency relationship binding them together would yield a Pareto optimum outcome only under two stringent and intuitive conditions. The first condition is that the provider would act in the best interest of the patient. The second is that the patient would comply with provider's treatment instructions. Violation of both or either of these conditions is quite common in actual treatment situations. Even when the provider is acting in the best interest of the patient, the patient may not follow treatment advice. There is also the possibility that the provider may be self-interested when treating the patient. Thus for a system of user charges to have a good chance of achieving Pareto-efficiency, it must be accompanied by government interventions that make it difficult to violate the efficiency conditions. In particular, the interventions should be aimed at passing the necessary health information to households and at regulating provider behaviour. The regulatory intervention for instance can take the form of a legal requirement for a periodic peer review of treatment procedures and practices followed by providers both in the private and public sector. Information from such reviews can be publicized to discourage unprofessional provider behaviour, including the widely known problem of
supplier-induced health care demand (see e.g., Grytten, D. Holst and P. Laake 1990).

An important uncertainty aspect of health services markets concerns the fact that people cannot tell if or when they will need particular types of medical care. For example, a need for a particular medical treatment (e.g., eye surgery) might arise when income is insufficient to meet the treatment expense. Further, even if income is sufficient to cover the cost of the required treatment, payment of this cost may ruin financial stability of the household. The uncertainty about future health states or treatment costs carries the risk of inability to afford medical treatment at a later date or of paying excessively large medical expenses even if treatment is afforded. This risk imposes a disutility on a household or its members because of the threat of suffering it carries. That is, households as well as the society dislike this risk – the threat to good health or to financial stability due to poor health. It is aversion to illness-related risk that motivates households to purchase health insurance. It would be very expensive for a risk-averse household to alone protect itself from welfare risks related to sickness for it would have to pay in advance all by itself an agreed cost of treating an illness. If, however, all households were to pool their risks together, the individual cost of risk coverage for each household would be reduced considerably. Each household would pay in advance only a small fee to cover a potential medical expense at any unknown period in a lifetime – an expense which without risk sharing would have been borne only by one household. Loosely speaking (i.e., ignoring the probability distribution of illness among households), risk pooling reduces the cost of risk coverage for each household for two basic reasons. First, all households contribute a fee in advance towards a fixed cost of treating an illness in each period over a lifetime so that the larger the number of households, the smaller the per capita fee. Second, given the natural illness incidence in the population, only a few of the total number of contributors ever need medical treatment in each period so that the fee income contributed is sufficient to cover the medical expense.

The foregoing discussion demonstrates that risk sharing enables households to purchase health insurance. Nonetheless, a mechanism for facilitating risk pooling among households may not exist, and in consequence, a market for health insurance may in turn not exist. In most urban settings in developed and developing countries alike, permanent wage employment provides the mechanism through which individuals can
voluntarily pool health related risks. The workplace facilitates easy identification of health risks faced by employees because, as a common place for all employees it reduces the cost of getting information on their characteristics, e.g., educational attainment and previous illnesses. Risk identification by a potential insurer – determination of the average probability that a wage earner in a given workplace will suffer a particular health problem over a given time period, say one year – facilitates, in turn, an assessment of the average cost at which the risk can be covered. This average cost forms the basis for computing premiums at which health insurance (the coverage of health risks) can be bought and sold. The degree of accuracy of this information determines whether health insurance markets will exist and if so the extent to which they will function efficiently. Since this information is difficult to obtain for workers in the urban informal sector (who may face different health risks from their counterparts in the formal sector), markets for these risks are typically missing. Formation and functioning of health insurance markets are also facilitated by the ease with which the premiums paid by workers are transmitted to insurers by the employer.

In rural settings, risk pooling among farmers is normally facilitated by cooperatives or other farmer associations. A village unit or a community-based organization may also play this role. These group modes of facilitating risk pooling are not as effective in helping a potential insurer identify health risks of individuals because group members do not have a common location so that the cost of getting information about their characteristics is quite high. The high cost of assessing people's health risks also increases the premiums that can be charged to cover a particular risk, thus reducing the number of people who may wish to pay the premiums. There is also the problem that group members in these types of associations do not have regular incomes which can be used to pay premiums. For these reasons, formal health insurance markets, including health care prepayment schemes in rural areas of developing countries are rare.

To stress the argument on uncertainty so far, it should be observed that an important consequence of uncertainty in health services market is the need for health insurance. However, non-existence of health insurance markets is pervasive in rural areas and urban informal sectors in developing countries because of the absence or imperfections of risk pooling mechanisms. Thus if health services are being financed through user
charges, government interventions are required to facilitate formation of health insurance schemes. However, two well-known insurance problems must be kept in mind: the moral hazard and adverse selection phenomena, both of which are extensively discussed in the general economics literature. [See Ellis and McGuire (1990), for instance, for a discussion of supply-side aspects of the moral hazard and adverse selection.] These phenomena arise because of asymmetric information in the health insurance market. As explained below, they are a problem because they are responsible for the commonly observed defects in the market for insurance and health services.

The moral hazard problem, is so-named because of the potential welfare loss associated with inappropriate personal behaviour (e.g., smoking) of an insurance holder that is unobservable by an insurance carrier. Since such behaviour is conducive to illness, it could lead to an over-use of medical care at existing premiums and eventually to higher premiums and to a limited insurance coverage for all. The adverse selection problem is so-named because of the adverse effect of unobservable characteristic (e.g., a chronic illness) of an insurance holder on the proportion of good risks among insurance buyers. People with such a characteristic select themselves into a risk pool more often than healthy individuals so that an insurance carrier ends up with an 'adverse selection' of customers. The consequence of this, as in the case of the moral hazard, is an over-use of health services at existing insurance premiums, a situation which eventually leads to higher premiums and to a limited risk coverage for all. Thus, the problem of moral hazard or adverse selection in insurance or any other market arises because of information hiding by an agent concerning his actions or characteristics that are of relevance to another agent (see Stiglitz 1995, p. 287). The hidden actions are endogenous, in the sense that the agent can influence them over the course of the insurance contract (e.g., heavy drinking or careless driving) whereas the hidden characteristics are exogenous, since the agent cannot influence them over the duration of the contract (e.g., previous history of mental illness or a tendency to suffer from allergies during the dry season).

Notice that insurance-financed medical care has the characteristic of a commons – a resource that is freely open to all for use in the sense that all insured patients can seek medical treatment free of charge. However, while the unit cost of medical care facing each insured patient at the time of use is zero, the social cost (the cost borne by the insureds as a group) is non-
zero so that rational behaviour by each patient leads to an over-use of medical care. Thus, even without the moral hazard problem (that is, even without a careless, unobservable personal behaviour that precipitates a need for medical care), the 'commons characteristic' of medical insurance provides an incentive for patients to over-use medical care. Insurance carriers dilute this incentive by imposing co-payments and deductibles on insured patients. These are the same measures that are also used to discourage people who hold medical insurance (and hence are entitled to receive medical care free of charge) from engaging in behaviour that might make them prone to illness. However, in situations where non-price rationing of health services is important (e.g., rationing by queues or travel time) these measures may be unnecessary. This is particularly the case in rural areas where patients typically incur high transport costs in form of travel time and fares to get to health facilities.

Notice that while co-payments and deductibles weaken the demand effect of the commons characteristic of insurance directly by restricting access to medical care, it reduces only indirectly, and probably only marginally, the demand effect due to the moral hazard attribute. These measures indirectly discourage moral hazard behaviour (behaviour that unnecessarily precipitates the need for medical treatment) by raising the cost of treatment, rather than by directly discouraging such behaviour. Direct discouragement of such behaviour by non-monetary instruments such as particular forms of institutions (see below) might be the correct measures however. This point can best be illustrated by an example. Medical insurance for treating malarial illnesses in a village, other things being equal, would lead to excessive number of visits to a local health facility due to both the commons and the moral hazard characteristics of insurance. The moral hazard problem in this case would arise because with insurance that guarantees free medical care, the villagers would likely shirk, for example, from destroying mosquito habitats, with a concomitant increase in malaria vectors. In this case, formation of village committees to destroy malaria habitats might be a stronger mechanism of dealing with the moral hazard problem than the imposition of co-payments for malaria treatments. Needless to say, the two measures together would be more effective than one or the other. Observe that the co-payments are used to tackle the moral hazard problem via the market mechanism, whereas the village committees are instituted to deal with the same problem via the use of rules and regulations to elicit or restrain a particular form of behaviour.
In summary, for a system of user charges in the health sector to function effectively, it must be complemented with a system of medical insurance schemes because of the uncertainty attribute of medical care. It is evident from the above discussion that because of informational imperfections in health insurance markets, the private sector cannot be relied upon to provide health insurance to the population: a whole range of health insurance schemes would necessarily have to be initiated or managed by the government. Since patients are not well-informed about medical care technology, government intervention is further required to control service quality, especially in the private sector where quality is likely to be compromised in pursuit of profit. The point here is to stress the need for a regulatory role of government in a situation where health services are financed via user charges, rather than to specify the form that the regulation should take.

The preceding are other reasons for government participation in the allocation of health services even when user charges are being used for that purpose, in addition to the usual equity and merit goods considerations.

3.4 Institutions, transactions costs and scale effects

In its standard form, the perfectly competitive market model, which, as has been argued in section 2, provides the theoretical rationale for user charges, assumes that economic agents engage in atomistic competition for scarce resources and that they interact and transact costlessly in the marketplace. This is the assumption we examine in the context of health services market. To the extent that it holds, the assumption of atomistic health care providers, who compete with each other for customers, should assure service quality and restrain unnecessary increases in user charges in health care markets. Disregarding the informational problems, this assumption has some descriptive relevance for the private health care subsector characterized by solo providers, as in urban areas of many developing countries. It further suggests that there is some threshold level of user charges that cannot be exceeded in small-scale health facilities in the public health sector without turning a significant number of patients to the private sub-sector or to traditional healers and self-remedies. Thus, the assumption of the competitive market model has some policy applications for the health sector. Several questions, however, need to be answered: first, what makes competition and interaction in the health care market or in any other market an orderly activity? Second, what are the efficiency
implications of organizational forms and sizes of health service providers and users? Third, what should be the role of the government in view of these implications?

The efficiency result of the market mechanism depends on orderly and smooth interaction among economic agents in the marketplace. The key point to emphasize here is that this interaction is facilitated by institutions. If imperfect institutions prevail in the health sector or institutions of a particular kind do not exist, use of user charges to ration health services among the population would do little to improve efficiency or to raise revenue – the often stated objectives of user fees (Creese 1991; Wouters 1993; Collins et al. 1996). It should be noted that this conclusion holds as well if non-health markets (to which health care markets are linked), are plagued by inappropriate institutional structures, an issue that is not considered here. We argue that public policy is needed to create or rectify institutions in the health sector to facilitate smooth functioning of user fees as a mechanism for rationing health care services. That is, government intervention is complementary to user charges in achieving economic efficiency in the allocation of health care resources. We start by defining institutions and related concepts.

Institutions are the rules of the game of a society, or, more formally, are humanly devised constraints that structure human interaction. They are composed of formal rules (statute law, common law, regulations), informal constraints (conventions, norms of behaviour and self-imposed codes of conduct) and the enforcement characteristics of both. Organizations are the players: groups of individuals bound by a common purpose to achieve objectives. They include political bodies (political parties, the senate, a city council, a regulatory agency); economic bodies (firms, trade unions, family farms, cooperatives); social bodies (churches, clubs, athletic associations); and educational bodies (schools, colleges, vocational training centers) (North 1995).

The game of society in the health sector is the 'process' of allocating scarce health care resources among the population. The aim of this game is to find an allocation that is socially efficient. Its players are the households (users of health services), the providers of the services (public, private, local authorities, civil-society; and so on), and the central government (which enacts and enforces the formal rules under which the game is played).
The society's resource allocation game in the health sector, as a strategic game, is for each player, a set of strategies and the associated payoffs. The strategies are rational plans or courses of action aimed at making best use of the available health care resources in a given institutional context. The attendant payoffs vary in form, depending on the needs of the players. The payoff from a particular strategy might be better health in the case of a household, higher revenue in the case of a health care provider or more political power in the event of the central government. To highlight the issues involved, we give below an informal description of the nature of the game of society in the health sector and its players (for a formal treatment of games of strategy in general see e.g., Friedman 1986).

Within the health sector, there are many players of different types (i.e., many households, health care providers and government bureaucracies). There are also many governments, in the sense that there always exists a potential government capable of replacing an ineffective one. The play of the game – the process of allocating health care resources – is occasioned by a particular state of nature, namely: a state of poor health. The institutions of society entitle each player with a given endowment of resources that can be devoted to improving health status. Each player responds to the 'condition of poor health' in the best way possible. Resource scarcity in the health sector, as in any other sector of the economy induces competition among player types in the design of response or survival strategies. For example, health facilities which provide services at least cost earn more revenue at prevailing user charges than the inefficient facilities and can thus fulfil their needs better than inefficient facilities so that each facility has an incentive to minimize the cost of service provision. Households which seek medical services from least cost providers can satisfy a greater variety of health care needs in addition to other human wants than households that are not as cost conscious in seeking medical care. And similarly for other types of players such as the public bureaucracies, social groups, civil society and governments.

Under certain assumptions (see also section 2), the strategies just noted would in aggregate, lead to a socially efficient allocation of health care resources. Four of these assumptions are a) existence of institutions to facilitate competition among market participants in the health sector; b) zero costs of transacting in health care markets so that the benefits from
trade in these markets are fully exploited; c) small and individualistic users and providers of health care who take the economic environment as given; and d) non-price mechanisms do not matter in the allocation of health services and hence health care prices convey sufficient information for economic efficiency. Since these assumptions hardly hold in the health sector, government intervention is required as a supplement to user charges in the allocation of health care resources.

3.4.1 Institutions

In the context of the health sector, institutions are humanly devised rules that structure interaction among users and providers of health care services. Needless to say, these rules may differ in type and degree from the rules that govern interaction in non-health sectors or across sectors. We focus attention however on institutions in the health sector. Because of asymmetry of information between users and providers of health services – an asymmetry that puts users at a disadvantage in judging service quality – provider entry into the health sector is tightly regulated by health professions and by government ministries of health. Before a health care provider can set up a medical practice, certification of qualification by a medical body is required. In addition, the provider must certify licensure regulations of the Government. For example, in some countries, an individual cannot get a licence for private practice while employed in the civil service. In other countries, private medical practice may be forbidden altogether or the law of the land may not recognize rights to private property. With regard to regulation of entry, the institutional structure might consist of rules for government licensing of private medical practice and for certification of competence to practice. Legal recognition and protection of property rights is an important element of the institutional structure. This structure would affect the number of private providers in the health sector and hence the degree of competition among providers. In turn, the size of the private sector should affect the degree of price competition between public and private health care providers and hence the average prices for health services in the two sectors.

Some elements of the institutional structure just noted – the certification rules, for instance – also affect the extent of competition among providers within the public health sector. Because of rigorous certification process for medical doctors and related personnel, only a few health facilities in the public sector can be staffed with this cadre of personnel. Thus, the range of health care providers of a given type available to patients in the public
sector is quite limited. In consequence, competition in health services provision is imperfect. It is important to note that the source of this imperfection is the prevailing institutional structure that affects the supply side of health services market. In this case, imposition of correct levels of user fees in the health sector (i.e., getting the prices right) without altering the underlying institutional structure would do little to improve the health of the population.

The issue of institutional structure and its effect on competition in the public sector is much deeper than it appears. Competition among providers in the public sector is not only affected by stringent certification rules that restrict the supply of qualified health personnel, but is also restrained by lack of motivation to compete. The structure of work incentives such as long service before promotion or assurance of security at the facility as a condition for additional medical supplies, which might have become the norm in the public health sector prior to the introduction of user charges, may not be conducive to competitive behaviour. That is, imposition of user charges by itself is not sufficient to provide incentive for competition. The government needs to establish a compatible or self-enforcing incentive structure as it introduces or endorses user charges in the public health sector if benefits of competition in service provision, as seen in the standard price theory are to be reaped. An example of an incentive structure that would spur competitive behaviour in service quality among providers in the public sector is an enactment of rules for sharing revenue from user fees between the health facility collecting the revenue and the central government. The higher the share of the fee-revenue kept by the collecting facility, and the greater its autonomy in spending the revenue, the greater is its incentive to maximize fee-income, e.g., by improving service quality to attract patients from other providers. The same behaviour can further be encouraged by enacting regulations under which public facilities can alter service fees (to attract patients or to cover costs of better quality) with a minimum of bureaucratic authorization from the central government. That is, a regulation that empowers health facilities to alter fees routinely, but within certain limits, without having to go through cumbersome approvals by the central government is a form of work incentive to health personnel, for it reduces the costs of responding to unforeseen events. Such a regulation should motivate workers to respond to events to which they would ordinarily have done nothing because of the costs involved, with possible adverse effects on performance.
It has been noted that retention of fees at the collecting facility is a form of work incentive. Even though the revenue retained at the collecting facility in the public sector cannot be put to private use, it can be spent on items that the management considers important in public service. It may thus improve work morale, an important factor in the fostering of competitive behaviour among health care providers in the public sector. Notice that it is not essential for an incentive structure to serve self-interest for it to spur competition. The crucial requirement to this end, is facilitation of achievement of a social objective which is important to health workers or to health facility managers, e.g., discharging a commitment to improve health status of a particular community (see e.g., Sen 1979). The retention of fee-income at the collecting facility could help achieve such an end, by encouraging health workers to improve the services they offer with a view of attracting particular types of clients (e.g., attracting mothers who would otherwise have gone to a traditional birth attendant for delivery to government hospitals).

3.4.2 Transactions costs

We begin by defining transactions costs. 'Transactions costs are the costs of specifying what is being exchanged and of enforcing the consequent agreements' (North 1994, p. 361). In the idealized Walrasian model of resource allocation, agents transact in the marketplace at zero cost because the information they need about what is being traded is assumed to be freely available or they are assumed to already possess such information. Moreover, the cost of enforcing an agent's agreement to buy or sell is a trivial one because time does not matter in concluding a transaction: once the price is agreed upon, the selling and buying occur instantaneously. However, in real world markets, especially in the market for health services, there can be substantial costs of acquiring and processing the information needed for an exchange; the costs of enforcing exchange agreements can also be quite high. For instance, patients need information both about their illnesses and about the quality of the available health services before they can make treatment decisions, i.e., decisions concerning medical care source and the quantity of medical services to use. The decision as to the source of treatment is typically made by the patient, whereas the decision concerning the type and quantity of medical treatment is made by a health care provider on behalf of the patient. In either case, however, the patient requires certain types of information before these decisions can be made. In deciding on treatment source, the patient needs information about the quality and cost of medical care.
offered by various providers. As to the decision concerning the type and quantity of medical care, both the patient and the provider need information about the nature and severity of illness, which is typically acquired through diagnostic tests. The acquisition and processing of these sorts of information require resource expenditure by the patient over and above the expenditure indicated by published fees for medical treatment. Further, there are potential costs of enforcing treatment agreements (the implicit agency relationship between the patient and the medical care provider) that might affect the choice of treatment source or type. Notice that unless the patient is able to meet the informational or enforcement costs associated with treatment, a transaction between the patient and the provider (an exchange of money for a service) may not take place. Thus, the ability of patients to meet informational and enforcement costs in health care markets affects the volume of transactions in these markets: the greater this ability, the larger the number of visits to health facilities over a given time period and conversely. In other words, because transactions costs have to be borne by patients, the actual demand for medical care is smaller than the quantity presumed under the posted levels of user charges. Government interventions, for example in form of disseminating information about type, quality and location of the available health services can reduce the transactions costs faced by patients and thus raise the rate of service utilization at existing fees. Government revision or tightening of enforcement of regulations governing claims of medical insurance benefits might also raise demand for medical care without changing the existing levels of user fees. Further, a classification of medical practitioners and health facilities that signals to households the types of services being offered at different levels of the health care system can serve the same purpose. Subsidized fees for diagnostic services are an additional way of reducing transactions costs faced by patients. Thus, selective interventions by the government can be used to reduce transactions costs in health care markets and complement a system of user charges in improving the health status of the population. The foregoing discussion suggests that introduction of user charges in government health facilities implies a diminished role for the government in the funding of curative health services but suggests also an expanded role in promoting the evolution of institutions under which user charges would allocate health care resources more efficiently. Thus, introduction of market-based reforms in the public health sector such as the implementation of user fees may increase or reduce the role of the government in this sector depending, among other things, on the extent of development of institutional structures
(e.g., health insurance regulations, public health law and practices, professional codes of behaviour) in a particular country.

3.4.3 Size effects

In the context of the health sector, the allocative efficiency result of the price mechanism depends partly on the assumption that households and health care providers are small so that none alone can influence the structure of economic incentives that govern the use of health care resources. This assumption is particularly problematic when considering effects of price reforms in the hospital sub-sector of the public health care system. Government hospitals in most developing countries are monopolies in the localities where they exist because of the large expense involved in building them. Introduction of user charges in these hospitals does not change their status as monopolies. Consequently, the inefficiency problems associated with monopolies apply to these hospitals as well so that imposition of user charges in and of itself cannot yield desired results in the health sector. Specifically, because of the monopoly status of hospitals in a given geographic area, the quantity of health services provided to the population would typically be smaller than the one that would prevail in a competitive situation. Or, equivalently at given levels of user charges, service quality would be lower than the quality that would be realized in an environment in which health facilities compete for patients. In view of the latter consideration there is need for government to design mechanisms for quality assurance in hospitals as user charges are introduced. However, since governments control the pricing of hospital services, user charges can be set so as to force hospitals, as monopolies, to provide the same level of services as in a competitive environment. Further, user charges can be set at levels that reflect any scale economies in service provision. The role of central government is important in enforcing the payment of user charges and in designing a system for accounting for revenue from fees. These tasks cannot be left to the management of individual health facilities.

The users of health services might be organized in form of trade unions, cooperatives or professional associations which could influence the setting of user fees in their favour relative to other social groups. Under such circumstances, user charges could be a major barrier to health care services by unorganized individuals in society. This is particularly the case in many developing countries, where urban hospital services are charged at the same rate as hospital services in rural areas even though transactions costs
are higher in rural areas (due to poor transportation to health facilities and the general lack of diagnostic equipment). In other words, the effective cost of medical care is lower in urban than in rural areas mainly because of the political power of organized health care users in urban areas which is used to influence health policies, e.g., the level of user charges. Since the sizes and organizational forms of health care consumers (such as trade unions or large urban corporations) influence the 'rules' under which health care resources are allocated in the population, the role of the market mechanism in allocating health care resources can be enhanced by specific public interventions such as exemptions of certain geographic regions or social groups from payment of fees or by subsidizing fees paid by particular social groups such as children or expectant mothers. Because of such political economy considerations, curative health services are best financed by a combination of user charges and revenue from general taxation. Further, the political economy issues in the allocation of health services among the population suggest that in developing countries, where the majority of the people suffer from infectious and parasitic diseases, government may need to increase its funding for curative services even as user charges are introduced in the public health care system if the health status of the general population is to be improved as rapidly as possible. Consumers and providers of health care services, when organized in large units affect not only the prices at which health services are available (by using their bargaining power) but they also by using their political power affect the political process that sets the rules under which health care resources in the health sector are rationed among the population. There is need, therefore, for government to continue funding health care in the public sector even after the introduction of fees to ensure that the health of unorganized sections of society is protected.

3.4.4 Non-price mechanisms

As noted earlier, when markets are perfect, the price system is an efficient mechanism for allocating resources. In that case, and other things being equal, a system of user charges is an efficient mechanism for allocating medical care among the population. However, in situations where markets for health services are imperfect or are non-existent, non-price factors such as directives or regulation, contracts, reputation, trust, loyalty and commitment are important in the allocation of health services. They may complement or replace the incentive function of the price system.
As a slight digression, it should now be noted that the incentive role of the price system works both on supply and demand sides of a health services market. An exogenously given system of user charges encourages patients to use medical care rationally because the charges signal, albeit imperfectly, the opportunity cost of resources patients expend on medical care. Similarly, the fees paid by patients are an incentive for medical care givers to be cost conscious in service provision because such behaviour increases the net revenue they receive from fees. However, when markets are imperfect, due for example to informational problems, the incentive structure of the price system may be misleading. On the demand side, low user charges at a particular health facility might be perceived by patients as a signal of poor service quality rather than as an indication of savings that can accrue to them if they were to seek medical care from that facility rather than from alternative sources. On the supply side, providers might incorrectly perceive low user charges as a form of subsidized prices rather than as an inability to provide quality services for which patients would be willing to pay higher prices. In both of these situations gains from exchange are not fully exploited because of information imperfections.

In an imperfect information environment, such as the one just described, non-price factors are crucially important in determining the behaviour of patients and providers in health care markets. To begin with, it should be noted that medical care is an *experience good*, the quality of which is observed only after consumption rather than a *search good*, a good whose quality can be determined before experience. To judge medical care quality, patients rely on their own experience in treatment and the trust they have in a provider, rather than on price information. They use the same method to make the appropriate treatment decisions. Further, reputation of health care providers (i.e., other people's experience in treatment) and the observed loyalty to a provider (frequency of return visits to a provider) are more important than the posted user charges in determining health care decisions. In consequence, providers respond to this situation by investing in patients' trust and loyalty to stimulate demand for their services. Investment in patients' trust also enhances provider reputation; it may take the form of additional personal attention to patients, provision of some services free of charge, a transitory improvement in service quality; and so on. Because of such investment, which is recouped by eventually charging higher fees (see below), providers may not be interested in correcting information imperfections in health care markets by disseminating health information, for example, to households. It has
been shown that in an imperfect information equilibrium, user charges in the *overall* health care system are higher than in a perfect information equilibrium (McGuire 1982). Under this circumstance, government regulation on provider behaviour, public dissemination of health information, and an increase in staff morale in public health facilities would improve the pattern of medical care use and provision.

It is important to note that the foregoing discussion proceeds under the assumption that effects of fees in the public sector on medical care demand in the private sector are being considered. Thus introduction of user charges in the public health sector raises the costs of medical care in that sector relative to costs in the private sector. Other things being equal, patients switch to private health facilities, a behavioural response which under imperfect information situation, could spur investments by private providers in patients' trust and loyalty. To recover this investment, providers eventually raise health care prices, without turning away patients because of the 'loyalty effect' on demand. Essentially, investments in patients trust coffer a monopoly power to a provider, thereby enabling him or her to charge the patronizing patients far above the marginal costs of service provision. The key point here is that any observed shift of patients to the private sector after an introduction of user charges in the public sector might not be due to better service quality in the private sector but rather to patients' trust in private providers.

Observe that trust reduces the uncertainty about medical care quality and therefore increases the *ex ante* benefit from treatment. However, the *ex post* benefit from treatment must surely be smaller, because trust exaggerates the provider's ability to treat an illness successfully. Patients' trust enables providers to receive rents for their services, i.e., to be paid fees which are over and above what they need to receive in order to remain in business. Thus, the resources invested in trust formation are misallocated. As already noted, this misallocation arises because of information asymmetry, and can be mitigated or avoided by informing the public about the types of medical services available in government as well as in private health facilities. Since such information is a public good, it is best financed by the government. Consequently, introduction of user charges in public health facilities implies an additional financial responsibility for the government. Here is an instance, as elsewhere in this essay, where introduction of a market-oriented reform calls forth an
additional role for the government in order for the reform to achieve its desired aims.

The revenue required to finance public health information can be obtained by reallocating additional funds to the health sector or by increasing the tax effort. It is important to note that in this instance, increased taxation aimed at funding the health sector is not a substitute for imposition of user charges for health care, as is often proposed in the health economics literature, but a fiscal measure intended to improve performance of a system of user charges (see below). Widespread imperfections in health care markets necessitate this public policy.

3.5 Coordination and regulation of actors in the health sector

3.5.1 Rationalizing coordination

The ultimate aim of a reform, such as the introduction of user charges in government health facilities is to improve the health status of the population. Needless to say, realization of this basic aim is conditional on the reform achieving intermediate objectives such as efficiency and equity in health service delivery. It has already been noted that efficiency (or efficiency-with-equity) in the health sector is an outcome of a complex societal game played by many actors within a particular institutional structure. It is important to emphasize that whether or not a reform such as the implementation of a system of user charges in government health facilities improves health status depends critically on the institutional structure of society. In addition to institutions that reduce uncertainty and transactions costs of economic agents in the health sector, for example, the laws on health insurance claims and the mechanisms for their enforcement, there are institutions which ensure that the agents as a group act in a manner that is consistent with the overall aim of the reform. Examples of such institutions include public statutes that govern the operations of health facilities in the public and the private health sectors. If the institutions that coordinate or regulate the behaviour of the various participants in the health sector towards a common goal do not exist at the time of the reform, its intended purpose would not be realized or would be realized only partially.

The great advantage of the Walrasian price system is its perfect coordination of activities of the numerous agents in an economy
characterized by the assumptions of perfect markets. Market clearing prices in such an economy convey to agents all the information they need to make concerning individual consumption and production decisions so that in aggregate, the best possible allocation of resources is achieved. Similarly, equilibrium user charges in a public health sector (also characterized by the assumptions) are sufficient for efficient resource allocation in that sector. However, as stressed at several points in this essay, the real world phenomena of externalities, information asymmetries, indivisibilities in consumption and production, among others, make individual responses to price signals in any sector, particularly the health sector, inefficient from an aggregate standpoint. That is, an efficient resource allocation decision at the individual units in the health sector is inefficient from the viewpoint of society. For this reason, there is need to coordinate or regulate decisions and activities of individual health care users and providers so as to ensure a socially desirable allocation of health care resources. Since the issue at hand is the enhancement of health status for the whole population, the coordination can be done only by a public authority, which has both the power to coerce as well as to use public funds to induce economic agents to act in social interest. That is, coordination or regulation of the health sector may be fiat-based or incentive-based.

Fiat-based coordination in the health sector (or any other sector) is in the spirit of central planning, in that individual economic agents are issued with performance directives by a central authority. This form of coordination is informationally very demanding on the part of the regulator, for the regulator is assumed to possess detailed information about preferences of health care users as well as the production technologies of health care providers. Since such information cannot be acquired by a single individual or agency, fiat-based regulation is less effective than incentive-based regulation as a mechanism for controlling economic activities. No attempt is made in this paper to survey the vast and technical literature in this area (see, for example, Laffont and Tirole [1993] for various approaches to regulation).

In a decentralized system of health service provision where health services are allocated by a system of user charges, incentive-based regulation is the appropriate mechanism for controlling behaviour of economic agents. In an incentive-based regulation, the aim is to provide inducements for sovereign health care users or providers, who from a social perspective
face incorrect market signals to act in social interest. We discuss only regulation of provider behaviour. The basic point in an incentive-based regulation is that health care providers are in a principal-agent relationship with a government regulator (e.g., a planning agency, a public commission, an inter-ministrial committee, or a government ministry). The relationship may be implicit (as for example, between a government hospital and the ministry of health) or explicit, as for instance, between a private hospital and the health ministry). In both cases, the hospital functions as an agent of the ministry of health – the principal in delivering health services to the population.

The aim of the regulator might be to control quality or prices (user charges) by using contractual or fiscal instruments to make the agent behave in a particular way. In the public health sector, quality control is typically the relevant domain of regulation because user charges at public health facilities are set by the central government. In the private health sector, however, prices (or user charges) are also subject to regulation because of the monopoly power of providers which tends to keep service fees far above marginal costs. We concentrate primarily on quality control of the services rather than on the control of the fees charged.

To begin, quality control is an issue in the health sector or in any other sector when an agent supplies goods or services to the government (or to the public on behalf of the government) in the context of imperfect information. In the health sector, the former is typically the case when private firms supply drugs and medical equipment to government health facilities. In the second case, health care providers give services to the public on behalf of the government, as for example, when missionary or private hospitals treat patients free of charge under cover of a government subsidy. These are important issues in the financing of government health services for two reasons. First, the quality of drugs and medical supplies purchased from the private sector by the revenue from user charges or from taxes determines the efficacy of the public health care system in dealing with the prevailing health conditions. Second, private health facilities might be more cost-effective in assuring quality of certain types of health services than government facilities so that health needs can be met more effectively by contracting out provision of such services to the private sector. In these cases, the regulatory issue concerns the design of an incentive scheme that would motivate providers to deliver services of sufficient quality at least cost. Two commonly used incentive schemes
(Laffont and Tirole 1993) are cost-plus-fixed-fee contract (cost-plus contract) and the fixed-price contract. Also, see Kutzin (1994) for examples of contracting in the health sector in developing countries.

3.5.2 The cost-plus contract in an implicit form

In the cost-plus contract, the regulator reimburses the provider the full cost of service provision and in addition pays the provider a fixed fee. This is a low powered incentive scheme because the provider has no inclination to deliver services at least cost since all costs are paid for by the regulator. This is essentially the type of incentive scheme available to managers of government health facilities (hospitals, health centers and dispensaries), before and after the introduction of user fees. The government meets the cost of medical care at its health facilities and then pays the health personnel a fixed fee in form of salaries. This is the case with government hospitals, which, after the introduction of fees are expected to generate their own revenue. It should be noted that under the cost-plus scheme, the incentive of the health personnel to improve service quality would be weakened if, in an attempt to reduce its share in health services expenditure, the government were to reduce the health budget in proportion to the revenue from user charges. Moreover, the motivation to improve service quality would still be weakened even if the size of the health budget were to remain the same, as for example when the revenue from user charges is committed to cover expenditure previously financed by the central government. This expenditure commitment gives rise to a performance disincentive because the revenue from user charges comes at the expense of the effort the health staff make to collect it, and yet it is not available for discretionary use.

Under the cost-plus contract, the government regulator can motivate health personnel to endeavour to increase service quality by using some proportional matching of the revenue raised via user charges with an additional allocation from central government, part of which can be used to finance 'discretionary' health activities. We consider health activities to be discretionary if they are initiated by the health personnel at a health facility in consultation with the community served by the facility independently of the government regulator. Notice that the matching arrangement is unworkable if the level of the health budget in real terms is held constant from year to year subsequent to the introduction of user charges. The matching proposal rests on the assumption that even after the introduction of user charges, budgetary allocations to health facilities continue to grow
each year at some constant rate thus permitting the matching to be financed by annual budgetary increments from the central government. The essential point for the viability of the matching mechanism is for the health budgetary allocation to continue to grow (even at a lower rate) after the introduction of user charges. This particular incentive system implies an exertion of a higher tax effort if the growth in health budgets cannot be sustained via a reallocation of the revenue generated by existing fiscal measures.

Thus, introduction of user charges at government health facilities might require complementary fiscal measures in order for the purpose for which they are introduced to be realized. The fiscal measures might be in form of earmarked taxes for health care or in form of more flexible and broadly based revenue collection methods such as presumptive tax techniques (see e.g., IMF 1996; Taube and Tadesse 1996). In particular, presumptive taxes – taxes based on presumed income of economic agents merit special consideration because they are typically targeted at sections of the population which are not already covered by the tax net (e.g., taxes levied on the basis of cost of business licences issued to persons not registered as tax-payers). Thus, presumptive taxes involve a widening of the tax base rather than an increase in the tax rate and so their efficiency loss is likely to be small. Further, the negative Laffer effect of such taxes on government revenue (the decline in tax revenue when a tax rate exceeds some threshold level) is likely to be unimportant because marginal tax burdens would not be affected and so taxable income should not decline. However, it can be argued that presumptive taxation for health care, which would be broadly based, even if feasible, should not be levied when user fees are also in effect because it amounts to a double payment by the population for medical care. This is the argument that underlies a frequent suggestion that tax revenue be used as a substitute for user charges in the financing of health services.

Presumptive taxes when levied in conjunction with user charges both constitute a double payment for health care. However, the double payment is necessary for two reasons. First, the tax revenue, when collected and earmarked for collective medical care, is a form of health insurance. Second, as argued previously, when medical care is financed by insurance, a system of moderate user charges serves as a mechanism both for curbing the moral hazard behaviour as well as frivolous demand.
Since the tax revenue is paid in advance of illness, use of this revenue to finance health care for all is redistributive in favour of the sick (persons with low stocks of human capital). Furthermore, tax revenue can be used to finance exemptions granted to the poor people who are unable to pay moderate fees that need to be levied at government health facilities for efficiency reasons.

Two characteristics of the health care financing strategy just described should be noted. First, both service provision and funding are in the hands of the government. Specifically, the government is the carrier of the medical insurance as well as the provider of medical services. The non-separability of these functions arises from non-existence of insurance markets in low income areas. The financing of government health services through user fees without the mediating mechanism of compulsory public medical insurance would be inefficient and inequitable. The fees rationalize medical care use, by mitigating inefficiencies due to moral hazard and the commons characteristics of insurance. The publicly financed insurance scheme helps smooth service use over time both by the rich and the poor. Notice that the problem of adverse selection does not arise here because health insurance is compulsory. The financing of government health services via user fees without a publicly mandated insurance system is equivalent to a fee-for-service system in the private sector without the mediation of private medical insurance. That, such a system would lead to immense inequalities in medical care consumption is straightforward to see.

The second characteristic of the foregoing strategy is that even though people have a common, publicly mandated medical insurance for treatment at government health facilities, they can choose to buy private insurance to finance medical care outside the public sector. In that case, the co-payments typically imposed by private insurers play the same role as that played by moderate user fees in public health facilities. However, in a low income area, only a small fraction of the population would afford to purchase private medical insurance. People with such insurance would still be required to pay taxes to finance publicly mandated insurance at government facilities for equity reasons, i.e., to help finance health care of the poor, since medical care is a merit good. Apart from this ethical reason, there is precautionary motive for such a requirement: publicly mandated insurance provides people with a fall-back mechanism for getting medical treatment in the event of inability to sustain private insurance.
Thus, in a low income area where medical insurance schemes are difficult to establish, health care financing via a combination of user charges and taxes promotes both efficiency and equity. We stress that to achieve efficiency and equity in health service delivery in a low-income area, use of both fiscal and price instruments is required, with emphasis initially being placed on fiscal instruments. Some form of direct intervention, e.g., in service provision and management might also be required. The need for these administrative interventions arises because of non-existence or imperfections of health care markets. The role of fiscal instruments and administrative interventions in health care provision and financing should gradually be reduced with economic growth, and with evolution of institutions conducive to market transactions.

3.5.3 The fixed-price contract

So far, focus has been on cost-plus contracts as coordination devices. We now turn to fixed-price mechanisms. In a fixed-price scheme, the government regulator pays an agent an all-inclusive fee for health services contracted. This is a high powered incentive scheme because the provider appropriates any cost savings that occur, and thus has an incentive to be efficient in service provision. Fixed-price contracts are particularly suited for health care provision by private providers or by autonomous public hospitals, where managers are free to use any cost savings as they wish. In particular, the government regulator can issue a fixed-price contract to a private or a missionary health care facility which has a cost advantage in the provision of certain services (e.g., family planning and immunization).

The private sector might further have a cost advantage in the provision of non-clinical services routinely provided to patients by a government health facility, e.g., laundry, cleaning, and catering services. In this case, the regulator can use revenue from user fees more effectively by having the non-clinical services in a government facility performed by the private sector under a variety of fixed-price contracts. However, certain informational asymmetry problems make fixed-price contracts difficult to design. In the cost-plus contract discussed in a previous subsection, it is easy for the government regulator to get accounting information about service provision technologies at public health facilities. Because the government has full access to the accounting records of its facilities, such information is straightforward to gather. However, in the case of fixed-price contracts, accounting information on which to base contract prices is
difficult to get from private providers. Private providers are under no obligation to truthfully reveal their cost data to the government, particularly when they can use that information to negotiate a better contract price. Hence there are potential rents to be earned from such a contract, a situation that creates a possibility for a cost-plus contract to be less costly (to the regulator) than a fixed-price contract. In consequence, a regulator could prefer a low-powered incentive contract to a high-powered one. We discuss below regulatory situations under which both features of the two forms of contract might be needed.

3.5.4 Linear and non-linear incentive contracts

Linear contracts are incentive schemes for which the share of cost of service provision between the agent and the regulator is some positive fraction between zero and unity. That is, such contracts combine features of cost-plus and fixed-price contracts. The amount of fee paid to the agent by the regulator is some linear function of the cost of service provision. Non-linear contracts in contrast, have certain ceilings on service cost that can be covered or reimbursed by the regulator. These contracts can as well involve a change in the rate of cost or profit sharing after the cost or profit has exceeded some threshold level; they tend to be piecewise linear, but real-world incentive contracts are often linear (see e.g., Laffont and Tirole 1993).

The regulator faces two problems in implementing incentive contracts. The first concerns gathering information about service provision technology of potential contractors, i.e., agents. Technology is assumed to be exogenous to health care providers. That is, each health facility is powerless to change the art of technology it can afford to use over the contract period. However, the technology can change exogenously during the contract period. Further, the information on provider-specific technology contains much noise. Nonetheless, it assists the regulator to discriminate among agent types in terms of differences in cost structures thereby mitigating the adverse selection problem, i.e., the problem of awarding contracts to high-cost providers. The proportion of high-cost providers among contract applicants would tend to be higher than that of the low-cost providers because they stand to earn greater rents from contracts. This potential gain comes from the fact the regulator treats high-cost providers as high ability (efficient) providers in determining their service fee. Further, the rate of cost sharing between them and the regulator is the same as the rate for the low-cost providers so that in effect, the regulator shoulders part of their
inefficiency. These two phenomena (a) the possibility of the high-cost providers successfully presenting themselves to the regulator as low-cost providers because of the information asymmetry about technology and (b) the treatment by the regulator of high-cost providers as low-cost providers in awarding contract fees because of the inability to discriminate between the two provider types, are the source of the adverse selection problem in the implementation of incentive contracts.

The second problem concerns obtaining information for making an inference as to the effort the providers are likely to make to reduce cost. The cost reducing effort is endogenous in the sense that its level is freely chosen by the provider, e.g., the amount of effort a hospital management exerts to control wastage and pilferage of variable inputs such as drugs and linen or to reduce the cost of overhead inputs such as water and electricity. This effort is assumed to be unobservable by the regulator. However, accounting information on internal cost control systems of potential providers can form a reasonable basis for awarding contracts to providers who are likely to exert a high effort to reduce cost. The key information for this purpose is information on the internal cost control mechanisms in all activities of a provider. Such information would give some idea as to the general level of cost reduction effort in an entire organization. If cost reduction effort extends in all activity areas of a provider, coverage of the cost for one activity area (e.g., immunization) by an incentive contract, is unlikely to affect cost reduction effort in that particular area. This might arise because of spillover effects or the indivisibilities of the cost reduction effort which would continue to be exerted in other activity areas. As in the case of adverse selection, the moral hazard problem, i.e., the problem of awarding contracts to providers with a low cost reduction effort (low-effort providers) has two sources. The first is the poor information held by the regulator about cost reduction efforts of potential contractors. The second source is the high cost of getting better information (i.e., the expense of monitoring cost reduction effort of providers). In consequence, 'low-effort' providers can successfully present themselves to the government regulator as 'high-effort' providers.

3.5.5 A digression on price regulation

The responsibilities of the government regulator also extend to price controls. In the health sector, this involves regulating the fees charged for various services in government health facilities, especially autonomous hospitals. The fees charged for professional services in the private sector
may also be subject to government regulation. The basic motivation for public regulation of prices is market imperfections and political economy considerations, such as subsidization of medical care consumption by the poor. The regulation however should be applied selectively, with the aim of targeting price benefits to particular social groups. For optimal mechanisms of price regulation in the public sector, see Laffont and Tirole (1993).

3.5.6 Coordination, decentralization and centralization

The aim of the foregoing paragraphs has been to clarify key analytical concepts, the understanding of which is a prerequisite for the design of effective mechanisms for coordinating the behaviour of the various players in a decentralized health care system towards a desired social outcome. The essential point is that in a centralized service provision system (including the financing system), the regulatory activity which coordinates the behaviour of diverse providers so as to achieve an optimal social outcome must be centralized. Without a central regulatory body to create and enforce rules under which decentralized health care providers and users operate, a market-based reform, such as the introduction of a system of user charges in the public health sector would do little to improve the health status of the population. From this follows the observation that 'effective decentralization', in the sense of the ability to achieve a specified social objective, involves both privatization of particular types of property rights and the collectivization of other types, or privatizing certain sections of a public health facility while leaving other sections in the hands of the government.

As an illustration, to increase incentives for efficiency under a system of user fees in the public health sector, after an introduction of such fees, some sections of tertiary hospitals might have to be 'privatized'. For example, some of the doctors' offices can be rented out to private practitioners, for general outpatient care and consultations, while retaining the right to use other offices in the hands of the government. This would enable the government to intervene directly in the provision of health services both for equity and for merit goods arguments. Such an arrangement might be more cost effective than an intervention via various contractual arrangements with private agents as described previously. Thus, privatization of a government hospital need not involve transferring the whole of it to private hands. In fact, the discussion in this section shows that such a move is unwarranted both for efficiency and equity.
reasons since in many countries, markets for medical and related services are either imperfect or absent.

It should also be noted that health service decentralization involves two key actions by the central government: (a) empowering local level authorities, i.e., municipalities and community organizations to raise funds to finance government health services; and (b) further enabling these entities to spend the funds they raise in accordance with local needs and to influence the management of government health facilities. If however both of these functions are in the hands of the central government, health service delivery is decentralized. We have argued that the real choice that policy makers face is not between one or the other of these polar forms of health care financing and management, but an optimal mixture of the two. Considerations of equity, economies of scale and scope, among others, justify a centralized system of health care provision and financing; considerations of efficiency, effective targeting of public health expenditure, among others, justify a decentralized system. Which of these systems will dominate in a particular country, will depend on the extent of communication and interaction between the central government and local authorities. The prevailing political structures, the strength of the state, and the nature of community level organizations are also important in this respect.

IV MEDICAL CARE DEMAND AND THE IMPLEMENTATION OF USER CHARGES IN DEVELOPING COUNTRIES

4.1 Issues and assumptions

During the period of market-oriented reforms in 1980s and early 1990s, the debate in developing countries as to whether these reforms should be implemented in the health sector centred on three issues: (a) the extent to which the reforms could improve efficiency in the public health sector; (b) the extent to which they could alleviate the severe budgetary constraints at government health facilities; and (c) their probable impact on equity in the utilization of medical care. At the time of this debate, public health facilities were grossly underfunded, as evidenced by their frequent inability to meet recurrent expenditures (Creese and Kutzin 1995). In consequence, many of the peripheral health facilities such as health centers and dispensaries operated without basic medical supplies and people
hardly used them even though their services were available free of charge. In contrast, overcrowding prevailed at district and provincial hospitals, where service quality was somewhat higher; congestion at these facilities however hindered effective use of the available health care resources.

The theory of consumer behaviour suggested that introduction of user charges for government health facilities would generate significant revenue to meet part of the recurrent health expenditure. As a result, it was argued that it would be possible to (a) improve the quality of health services; (b) increase the rate of utilization of the government health facilities, and (c) improve equity in health care delivery using fee income. As noted in previous sections, the same theory suggested that introduction of user charges would rationalize the use of the available medical services. This would make it possible to increase service provision without an additional allocation to health ministries from the central government – a very attractive idea in a period of growing government budgets such as the 1980s. It was also argued that if part of the expenditure on curative health services could be covered by revenue from user charges, it would be possible to allocate proportionately more of the health budget to preventive health services – a more efficient allocation pattern since it is more cost-effective to prevent than to treat an illness.

The prediction of the above favourable effects of user charges – on revenue and service utilization – was based on the assumption that health services were inelastic with respect to user charges. Under this assumption, a modest charge on government health services would increase revenue without a severe adverse effect on their utilization. Consequently, a substantial amount of effort was devoted to estimation of price elasticities of demand for medical care [see e.g., Akin et al., (1985; 1986); Gertler et al. 1987; Dor et al. 1987; Schwartz, et al. 1988; Gertler and van der Gaag 1990]. From a theoretical standpoint, acquisition of this information was crucial prior to the implementation of user charges. The prediction of favourable effects of user charges was based on two other assumptions, namely: (a) the revenue from user charges would be used to improve service quality, with only a small time lag; and that the improvement would have a positive effect on health service demand – strong enough to outweigh the negative effect of user charges; (b) any households unable to afford health services at prevailing user fees would be exempted from the payment of fees. It was further believed that if health care demand were
inelastic with respect to user charges, the number of households unable to pay fees would be small and manageable.

There was also an implicit assumption that public announcement of fees by the government was the same thing as their implementation. The effort and expenditure required to educate the public about the fees, to train the health personnel in the collection and management of fee income, and to design and implement new management systems were either ignored or assumed to be negligible.

Introduction of user charges in developing countries, especially in Sub-Saharan Africa in the late 1980s and early 1990s revealed that all the above assumptions were incorrect. Specifically, the number of visits to government health facilities in many countries fell substantially despite the modest level of the fees implemented, and despite the inelastic health care demand response to these fees. Quality improvements in government health facilities generally did not materialize (see Creese and Kutzin 1995). Exemptions from user charges were declared with ease but proved to be very hard to implement because of difficulties in identifying persons unable to pay (see e.g., Gilson 1988; Huber 1993). In consequence, user fee reforms experienced temporary reversals in some countries (see e.g., Collins et al. 1996). We examine each of these issues in turn.

### 4.2 Price elasticity of the demand for medical care

The most well known and widely discussed econometric estimates of price responsiveness of health care demand for a developing country were reported in the mid 1980s for the Philippines, which were quickly followed by a number of similar studies (see e.g., Akin et al. 1985, 1986; Dor et al. 1987; Ellis 1987; Gertler et al. 1987; Schwartz et al. 1988). The Philippines study showed that the elasticity of medical care demand with respect to user charges was substantially less than unity. That is, other things being equal, a small increase in user charges would reduce medical care demand by very little. This result was derived with great care from a detailed household data set. Consequently, price inelastic health care demands were used in support of the policy of charging fees for health services in developing countries (see Akin, Birdsall and de Ferranti 1987; Griffin 1988). Further work on health care demand – stimulated by the Philippines study – showed that health care demand was highly responsive to user charges, especially at the lower end of the income distribution (see
in particular, Gertler and van der Gaag 1990). This latter study cautioned against a large drop in service utilization by the poor that could result following the implementation of even modest user fees in government health facilities. Nonetheless, user charges were implemented in many developing countries, especially in Africa, starting from the late 1980s, mainly on grounds that they would improve efficiency and sometimes equity (see Sauerborn, et al. 1994). In a recent and important paper -- because of its novel estimation of cross-price elasticities, Bolduc, Lacroix and Muller (1996), show that substitution effects of fees were not adequately addressed in earlier work on health care demand and this might have led to inappropriate health care financing policies. In particular, they show that the commonly made assumption of constant cross-price elasticities among providers may not be valid. We examine below the policy relevance of price elasticities of demand in the implementation of user charges.

4.2.1 Price inelastic demand

If health care demand is price inelastic, other things being equal, a small increase in user charges would increase revenue, with only a slight reduction in attendance at health facilities. Specifically, the proportional decline in attendance is smaller than the proportional increase in user charges. Thus, when health care demands are highly price inelastic, a system of user charges is a suitable mechanism for raising revenue for government health facilities. This policy conclusion however rests on two assumptions. First, a fairly high level of charges already exists in the public health care system. That is, user fees are not being introduced for the first time or being raised from a very low base. Second, other basic necessities, typically merit goods, e.g., education and clean water are not being sacrificed in order to maintain health care consumption at previous levels following an increase in user charges. We elaborate only on the first of these assumptions.

If user charges are being raised from a low base, a small increase in user fees can evoke a large percentage fall in health service utilization, despite a small price elasticity. As an example, suppose the existing user charge in the public health sector is FIM (Finnish Mark) 1.00, and the price elasticity of demand is 0.15. A small absolute increase in user charge from FIM 1.00 to FIM 3.00 (an increase of 200 percent) would lead to a 30 percent fall in health service utilization in that sector – a substantial decline. However, if the existing user charge were FIM 5.00, the same absolute increase in user
charge (a 40 percentage increase), would lead to only a 6 percent fall in health care use. This basic point was overlooked when user charges were being implemented in several developing countries, especially in Africa. As a result, large unanticipated percentage drops in health service utilization occurred in countries where user fees were raised from almost nothing to some modest levels (see e.g., Yoder 1989; Collins, et al. 1995; Kutzin 1994; Creese and Kutzin 1995; Appleton et al. 1996; Mwabu and Wang'ombe, forthcoming).

4.2.2 Price elastic demand

A price elastic demand does not necessarily imply that user charges should not be implemented in the public health facilities because of the large fall in service utilization that they would engender. The reason is that people might shift to cheaper alternative modes of treatment which are as effective as government health facilities in dealing with illnesses (see e.g., Bolduc, et al. 1996). In that event, social welfare would be higher than the welfare which would have prevailed had user charges been imposed in a situation of inelastic demands. The basic assumption that underlies the argument against the introduction of user fees in the public health sector when demand is price elastic is that there exists no suitable alternatives to government health services. This is a reasonable assumption for the rural poor who would likely shift to traditional or home remedies after the introduction of fees. In fact, a large proportion of the rural population shifted to low quality facilities (dispensaries) after the introduction of user fees in Kenyan government hospitals and health centers in December 1989 (Collins et al. 1996). Moreover, a large fall in attendance at government health facilities, mostly used by the poor, suggested that persons not using the public health care sector had turned to the informal sector remedies, including traditional medicine.

The assumption that the poor have very limited alternatives to effective health care outside government health facilities is a realistic description of the situation in low-income countries especially in Africa. Under this circumstance, a high price elasticity of demand for medical care implies that a modest increase in fees in the public health sector would shift people to unacceptably low quality medical treatments. Further, introduction of the same magnitude of fees in the public facilities for the first time, or equivalently, raising them from a very low base would have similar demand effects even if the price elasticity of demand is low [see Reddy and Vandemoortele (1996) for a recent policy discussion of price
elasticities demand for medical care]. Thus, high or low price elasticities in themselves do not tell the policy makers the welfare consequences of raising or introducing fees.

In summary, information on the degree of responsiveness of medical care demand to changes in user fees is not in itself a good indicator as to whether user charges should be implemented or adjusted. Information about levels of any existing fees, the nature and availability of alternative sources of medical care as well as the trade-offs that people might be forced to make after a policy change is as important as the information about price elasticity.

4.3 Medical care quality

The primary reason for introducing user charges in the public health sector was to improve service quality. In the early days of health care financing reform, it was assumed that the time lag between improvement in service quality and the introduction of fees would be a relatively short period. This however was not generally the case [see Litvack and Bodart (1993) for an exception]. Many health facilities in public health sector were not able to improve service quality after the introduction of fees for several reasons. First, it took a long time for facilities to get approval from the Government to spend the revenue they raised from user charges. Second, because of certain regulatory constraints, the facilities could not use public insurance funds to provide better services. Third, it took considerable time to set up financial management systems at health facilities, a prerequisite for an effective use of the fee revenue retained at the health facilities (Ellis 1987; Collins et al. 1996). The quality of health services improved only in a limited number of cases in Sub-Saharan Africa following the implementation of user charges (see e.g., Wouters 1991; Creese and Kutzin 1995).

4.4 Fee implementation procedures and health service management

Experience with user fee implementation in Kenya showed that its success depended on three key procedures: educating the general public about the fees; training the health personnel in the collection and management of the fee revenue; and designing and implementing an effective service management at all levels of the public health care delivery system (see Shaw and Elmendorf 1993; Mwabu, et al. 1995; Collins et al. 1996). All
these three procedures are critical in the performance of a new system of user charges as health service rationing devices. In the case of Kenya, user charges were suspended on September 1, 1990, barely nine months after their introduction, mainly because the above procedures were not followed. The fees were first introduced as *registration fees* on December 1, 1989, without preparing the public about their nature or purpose and without establishing the necessary financial management and control systems at the health facilities. The first problem subsequent to the announcement of fees by the government was their rejection, as a matter of principle, by the general public and by organized labour because people had been used to free medical care for twenty-five years. The second immediate problem was that people did not know the amount of money they were expected to pay at the health facilities as they had not been informed of the fees introduced at various levels of the government health care system. As a result, patients often arrived at government health facilities without the required *registration fees* – these were so-named because they were paid to register for outpatient treatment, irrespective of whether or not treatment was subsequently received. (A price system works because people know in advance what they are expected to pay to obtain a good or a service). The third and related problem was that the health personnel in remote clinics did not know what to do with the fee revenue they had collected, as they did not have the authority to spend it immediately, and did not have access to banking facilities. This problem contributed in part to their inability to improve service quality. Patients often paid registration fees, but did not get medical treatment because drugs were not available at health facilities. This fact in particular, engendered great public dissatisfaction with the fee system, and contributed significantly to its suspension by the government because of the adverse press publicity it received (Collins et al. 1996). It should further be noted that lack of effective service management at health facilities was perhaps an even greater obstacle to quality improvement than the inefficient financial system.

A cautious reintroduction of user charges, in which the above problems were rectified, led to public acceptance of user charges, and to some improvement in service quality (see below). In reintroducing the fees, apart from the public information campaign about them, the government devoted attention to training of health staff in the collection, recording and banking of the fee revenue. Patients paid the new fees after receiving medical treatment, an arrangement that greatly increased their acceptability.
Two managerial innovations were introduced during the second phase of user charges. The first innovation concerned creation of a central unit within the Ministry of Health to implement the new user fees. The unit was responsible for the actual implementation of the new system of fees. Its tasks included setting fees, initiating changes in national insurance regulations, and training health personnel in the new system of financial management. A second innovation involved establishing District Health Management Boards so as to decentralize authorization of expenditure of the revenue retained at the health facilities and to create representation of the community in the management of public health services.

The above changes were critical in getting the public to accept fees, and but they helped only marginally in improving service quality. Collins et al. (1996) note that:

Improvements in quality have not been consistent, partly due to poor use of funds, but also due to the inability of Treasury allocations to keep pace with rising costs, resulting in cost-sharing revenue being increasingly used to cover the shortfall in areas such as water, electricity, and patient food. Some hospitals have managed to improve quality by paying attention to low-cost quality improvements that matter most to patients, such as waiting time, cleanliness, and staff efficiency and attitude.

The main lesson from the Kenyan experience with user charges is that an effective implementation of a market-based health care financing reform requires much more than changes in price or monetary variables. Non-price changes such as regulatory and managerial changes are integral parts of the reform process. Moreover, even after successful introduction of user charges – in terms of getting these in place physically – their impact on health status depends on many other factors such as staff morale and their ability to allocate the revenue from user charges efficiently. The political economy of budgetary allocations to health facilities is also important in influencing the health outcome of the fee reform. This is because it affects the nature and availability of tax-financed health care inputs, which often complement the inputs financed via user fees. These non-price and non-economic factors were not given as much consideration as the price variables during the debate on user charges in mid 1980s and in the early 1990s. They were incorrectly assumed not to have an important bearing
either in the implementation or in the functioning of a system of user charges.

V THE HEALTH CARE FINANCING GAME

5.1 Preliminaries and the nature of the game

The policy discussion on key players in the financing of health services in developing countries has historically, and especially over the past ten years, focused on three actors: the private sector (including church and non-governmental organizations), the government (including local authorities) and the community (the collectivity of households). That is, the strategic game of society in the financing of health services has been assumed to involve three players. An additional player in this game, whose motivation and strategies have not been rigorously considered in the past, is civil-society — defined as a whole range of non-profit, voluntary associations in-between the household and the state (see e.g., White 1994).

As before, in addition to its players, the society's game is characterized by its rules, the strategies of the players, and the payoffs of the strategies. In this particular case, the aim of the game is to evolve a health care financing system that is non-exclusionary in service provision, i.e., a system that makes health services available to everyone in the population. The health care financing game is occasioned by a random occurrence of illness — poor health status. We assume that poor health is inflicted to the population by an 'invisible hand' of nature; thus, nature is the other player in the game.

The rules of the game consist of the institutional structure of society as defined previously. The payoffs sought by each player depend on the player's motive in the game. We assume throughout in the analysis that the necessary institutions for the play of the game by its participants exist; otherwise the description of the game is a parable rather than a representation of reality (see North 1995).

The type of motivation in the play is the distinguishing characteristic of the participants. The motive of the civil-society is assumed to be optimization of the welfare of its members without violating some ethical values of society. Thus, by assumption, the objectives pursued by civil-society
organizations are broadly shared by other players. In particular, the behaviour of civil society agents is strongly motivated by a desire to promote some social or ethic end. Thus, by a process of self-selection or comparative advantage, civil-society organizations come to be characterized by membership of individuals whose self-interest is strongly moderated by social and ethical concerns.

The above point requires comment because of its neglect in the mainstream microeconomic analysis. Adam Smith, in his *Theory of Moral Sentiments* (1759), strongly stressed the unselfish nature of human beings. However, economists have over the past two hundred years, largely ignored this part of Adam Smith's contribution to our understanding of human nature, preferring instead to adopt for their analysis, his concept of self-interest (Smith, 1776) as the cardinal force behind human behaviour. Bergstrom (1996, p. 1904) cites the following passages from the *Theory of Moral Sentiments* in support of unselfish as well as selfish nature of people.

How selfish soever, man may be supposed, there are evidently some principles in nature that interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it except the pleasure of seeing it.

Man feels his own pleasures and pains more sensibly than those of other people. The former are the original sensations; the latter, the reflected or sympathetic images of these sensations. After himself, the members of his own family, those who usually live in the same house with him, his parents, his brothers and his sisters are naturally the objects of his warmest affection ... his sympathy with them ... approaches, nearer in short, what he feels for himself. This sympathy too, and the affections that are founded on it, are directed by nature more strongly toward his children than toward his parents ...

The children of brothers and sisters are naturally connected by friendship, which after separating into different families, continues to take place between their parents... The children of cousins, being still less connected, are of still less importance to one another; and the affection gradually diminishes as the relation grows more remote (Bergstrom 1996: 1904-5).
The above passages show that human behaviour can be motivated by both selfish and unselfish reasons. For some economic agents, the selfish motive may be the dominant reason for engaging in their chosen activities, whereas for others, altruistic motivation is the main reason. We assume that the selfish motive of members of civil-society is strongly tempered by altruism, and in consequence, civil-society might help to internalize externalities of individual behaviour.

Health care financing as a strategic game of society is here conceptualized as a two-person, non-cooperative repeated game, played between nature, and a coalition of other players (see below). A profile of health care financing strategies by the coalition yields certain payoffs, expressed in form of improvement in health status of the population, given the strategy profile of nature. The strategy profile of nature consists of some probability distribution of illness or health needs in the population, which is unknown by the coalition prior to the playing of the game. Hence the non-cooperative nature of the game, for there is no mechanism by which the coalition and nature can come to a binding agreement as to what illness probabilities should be in force before the play takes place.

5.2 The solution concepts

The aim of the coalition is to evolve over time, health care financing strategies that can be rationalized as best – from the standpoint of health payoffs, given the actual and potential strategies of nature. (Notice that nature's response to the strategies of the coalition is another probability distribution of illnesses). The scope of the paper does not permit an elaboration of the extensive and normal forms of this game, which, as already noted is assumed to be played with repetition. Moreover, the best response strategies of the coalition at each stage of the game are not rational, in the sense of optimizing some single-valued functions under conditions of perfect knowledge. Rather, the best health care financing strategy profile of the coalition, given the strategy of nature, is the profile that is rationalizable. That is, a set of strategies (consisting, for example, of a system of user charges and tax-financed subsidies), which can be justified, rationalized, or shown to be reasonable on the basis of some criteria. For a distinction between rational and rationalizable strategies, see Pearce (1984) and Bernheim (1984).
A strategy profile consisting of user charges at government health facilities, tax-financed health care subsidies at private clinics, and exemptions from user charges for certain social groups can be rationalized or justified on grounds that it guarantees universality in the provision of medical care. In other words, no other set of policy instruments would dominate this particular set of strategies given the social criterion against which performance is being evaluated. In contrast to the solution concept of rationality in common knowledge games, as for example games with Nash Equilibria, the solution concept of rationalizability for games played under conditions of imperfect information requires only that the response strategies (and player beliefs on which they are based) be internally consistent. Further, the notion of equilibrium is not critical in solving such games because they do not have in-built equilibrating mechanism. Bernheim (1984) in his seminal paper, explains why the rationality assumption is not the relevant solution concept in games played under imperfect knowledge.

Since agents select ... strategies in ignorance of others' choices, they cannot optimize subject to actual selections of their opponents. As no sensible dynamic can eradicate this ignorance, we must justify any theory of strategic choice in terms of what it implies about the internal consistency of beliefs held by each player. If we begin to think of rationality in terms of internal consistency, it is obvious that players are not ordinarily compelled by deductive logic to select their Nash strategies (Bernheim 1984, p. 1009).

Rationalizability as a concept for finding solutions for strategic games is more realistic than the concept of instrumental rationality. In the latter, the solution for the game is sought in a set of single values of well-behaved payoff functions of players, playing under conditions of common knowledge [see Osborne and Rubinstein (1994)]. In the case of rationalizability, the solution is sought in the consistency of strategies in a situation of imperfect knowledge, which may be asymmetrically held among players. In the present case, strategies of the coalition are internally consistent if in application, no strategy contradicts the other as a best response to the strategy profile of nature. As Hurwicz (1953) states, internal consistency here implies 'a set of mutually undominated imputations' or strategies.
Rationalizability is not only a powerful analytical tool, but also a very useful policy device. It is a formalization of the nature of the everyday interaction among participants in an economy. Consider for instance, the case of policy makers in the health sector. To the extent that they are efficient in the design of health care financing strategies, policy makers critically examine these strategies for internal consistency, from a perspective of some logical premises, social values or goals. That is, they try to see if their policies or strategies are rationalizable. Policy research that deals with computations of single values, such as the price elasticity of medical care, or the probability of cure conditional on receiving a given treatment, is nothing more than an input into the process of establishing rationalizability of particular social policies or strategies. A rationalizable policy is easier to justify or explain to the public than one which is not. In turn, its implementation is less problematic, and can be sustained long enough to maximize the chances of it yielding the desired outcome.

The concept of rationalizability can further be used to illustrate a solution to the problem of designing an effective system of user charges for public health facilities. The relevant question to ask in this context is not the unresolvable question of whether the system designed by policymakers is rational, in the sense of having optimized some social objective function, but whether the system is rationalizable from the standpoint of some social evaluation criteria? If this sort of question had been asked prior to the implementation of user charges in developing countries or in countries in transition from central planning, many implementation problems such as suspension of fees or traumatic shock therapies might have been avoided. For instance, in the case of Kenya discussed previously, it would have been evident that the registration fees were exclusionary, since there was no exemption system in place. In consequence, the fees were inconsistent with the social policy of equity in health care provision. The registration fees were also inconsistent with the stated government policy of providing quality care since they were designed to be paid irrespective of whether or not service had been received. The lesson to be drawn here is that even though the system of user charges in Kenya had some very obvious defects, it was nonetheless implemented because appropriate analytical concepts were not used by researchers or policymakers to carefully examine its consequences or its nature. In consequence, the health care financing proposals recommended for implementation were not critically evaluated against alternative proposals.
5.4 The coalition of players and the special attributes of civil-society

5.4.1 The coalition, motivation for its formation and its characteristics

The coalition of other players, apart from nature, consists of the government, the households, the private sector and the 'civil-society'. Following White (1994), we define civil society as

an intermediate associational realm between state and family populated by organizations which are separate from the state and are formed voluntarily by members of society to protect or extend their interests or values.

Note that some organizations in this realm may not be civil; to that end, civil society has a negative externality. This issue however is outside the scope of the paper. There is need now to explain why the coalition of players forms. The players, other than nature, form a coalition for four reasons. First, the coalition avoids duplication of strategies against nature, for strategies cost resources to design and implement. Second, it permits exploitation of comparative advantage of all players in the design and implementation of strategies. Players design and implement strategies independently, but under the coordination of one of the players – the government. The government creates and enforces institutions that structure the behaviour of the coalition members. Third, the risk of an ineffective response to nature's strategy profile is smaller when all players act together (in coordination) than when they act individually (without communication). This is because of the simple fact that a solution to a problem is more likely to be found the greater the effort devoted to finding it (i.e., the larger the number of people involved in solving it). Coalitions are therefore formed to minimize the risk of inefficiency or ineffectiveness in response strategies. The assumption here is that all players are individually risk averse. This naturally leads players to desire a coalition, which they certainly form if other conditions exist, e.g., a mechanism for collective interaction or communication, such as a public radio or some other medium of mass communication. Finally, the coalition is formed to exploit scale economies as well as to internalize any externalities in the design of response strategies.

The civil society is particularly suited to deal with externalities because of the assumption that by self-selection, its members are persons interested in engaging in activities that promote social well-being. In order to fruitfully
evaluate the contribution of the civil society in the design of health policy, e.g., a system of user charges for public health facilities, there is need to first clarify its dual role as a user and a provider of health services. Associations such as welfare groups, with membership consisting of people without medical training are, like other economic agents, consumers as well as providers of health care. Such groups and their members use health services from modern health sector facilities such as the government hospitals as well as from other providers. It is important to note that the health services from these sources are obtained via the market mechanism; we elaborate on this below.

5.4.2 Civil society entities as health care consumers and the market mechanism

The health services used by group members are obtained at a fee and are provided by entities different from consumers. The payment of this fee by a member of a group, is much easier than its payment by an economically equivalent individual without a group affiliation for three reasons. First, a member of a group can borrow the funds required to pay for medical care from other members or from the group itself; thus group membership enlarges credit opportunities of an individual. Second, since a group is a large consumer relative to a single individual, it has the power to negotiate lower fees for its members with medical care providers. That is, group membership confers to an individual, benefits of scale economies in the financing of medical care. Third, the group is a mechanism for facilitating risk sharing by its members, with regard to uncertainty as to the ability to individually meet medical care costs. For this reason, group membership increases an individual's opportunity to enjoy benefits of medical insurance. Further, conditional on availability of medical insurance, group membership facilitates negotiations for favourable insurance premiums for it is cheaper for an insurance carrier to rate health risks of a group than the risks of an individual. Thus, on average, people with group medical insurance spend less on medical care than people with personal medical insurance. Briefly, membership into civil society associations lightens the burden of financing medical care primarily because it reduces the cost of transacting in health services markets.
5.4.3 Civil society entities as health care providers in market and nonmarket settings

Non-profit, voluntary associations are also providers of health care. The obvious cases here are the missionary or church medical associations. The distinguishing characteristic of such associations is their medical expertise. These organizations typically finance the subsidized medical care they offer to the poor through donations from their members, as well as from other people who share in their social concerns, such as the care for the poor. Existence of these organizations mobilizes funds from a wide spectrum of the actors, including the central government to finance health care of the poor. These civil society agents can also be used to accomplish altruistic or ethical objectives in the health sector. To that end, they are mechanisms for moderating selfish motives in the interest of social good. Note that the price system makes this moderation possible because persons unable or unwilling to participate in health care provision, say to the poor, can finance its provision. That is, the exchange system coordinates interests of health care providers and of people who may want to finance its provision to particular social groups -- groups in which providers may also be interested, such as the aged or the school children.

As can be seen from the foregoing, health service provision by civil society associations with strong expertise in medical care is extended to both members and non-members of such associations through the market mechanism. Provision of this service is financed by members as well as by non-members. In contrast, civil society associations without medical expertise provide non-technical medical care primarily to their members in non-market settings. Examples of such associations include self-help groups in rural villages and non-medical professional associations in urban areas. Members of these associations can, and often do exchange personal care in the event of illness. In this case, health care provision and financing are necessarily undertaken by the same person. Personal care, such as talking to patients, preparing meals for them and helping them in other activities of daily living, is financed by voluntary time contribution of the person giving the care. In terms of its contribution to the improvement in health status, personal care in non-market settings is not less important than the care provided and financed through the market system. Moreover, if expressed in monetary terms, the cost of this type of care can be enormous. Personal health care of the type just described is provided and financed by a large segment of civil society; however, it is rarely recognized as part of a nation's health care. From an econometric
perspective, the health care provided in non-market settings is an unobservable component of the national health care, and possibly, a very large error term of the measured component. Recognition of civil society as an important health care giver eliminates a substantial portion of this error term, and in consequence corrects the overstatement of the contribution of formal national health care systems to observed improvements in health status.

VI SUMMARY AND CONCLUSIONS

6.1 Summary

This paper has reviewed the economic theory and the institutional assumptions that provide the rationale for using fees to finance health services in low-income countries. The case in favour of user fees for health care in these countries, rests on powerful efficiency results of the price system as a resource allocation device. Under some very stringent assumptions, a system of user fees would allocate health care resources efficiently, with any equity concerns being addressed by simple fiscal instruments. No other resource allocation mechanism can dominate it. Not surprisingly, the case against user fees, turns on the stringency and the unrealistic nature of the assumptions that underlie the efficiency results. The most contentious being the assumption of perfect markets, and by implication the assumption of existence of institutions that facilitate the formation and functioning of markets. Also contentious is the selfish motive that is assumed to undergird much of human behaviour. The consequences of these assumptions for health policy have been explored at some depth in the paper. The policy value of price theory is to display the abstract conditions under which maximum efficiency in resource use in the economy or in a particular sector of it can be attained. Without these assumptions the performance of the price system as a resource allocation tool cannot be fully explored. The challenge in policymaking in respect of the price system is twofold. The first and obvious task is to determine whether the conditions under which the price system (a system of user fees) would work exist before it is applied. The second task, and the most daunting, is to design and implement reforms that would facilitate the evolution of the missing conditions. For example, there is no point in relying primarily or exclusively on the price system to allocate goods and services among the population if markets for many of the goods and
services do not exist or are defective. In the particular case of the health sector, the application of user fees faces many challenges because of missing and imperfect markets.

The application of a system of user charges presumes the existence of perfect markets for health care. However, price theory says nothing about formation of markets. It only shows how prices work if markets exist; it has no policy prescriptions to offer in this area (North 1994). It is here that the New Institutional Economics charts the way for public policy. It suggests that creation of certain institutions would facilitate the emergence of new markets or improve the performance of the existing ones. For example, enactment and/or enforcement of insurance laws would create favourable conditions for formation or better functioning of health insurance markets, for the rules of operation in such uncertain markets would be clear. That is, costs of transacting in health insurance markets would be reduced, inducing an increase in the quantity of insurance services supplied and demanded. This example indicates that in order for the implementation of user charges for health care to be effective, it should be undertaken in conjunction with certain institutional changes.

Correct institutions are not the only prerequisites for efficient operation of health markets or any other markets. The type of market information possessed by economic agents also matters. Information asymmetries in health care markets are an obstacle to proper functioning of fees: patients may pay fees for an ineffective medical service. As it happens, this asymmetry cannot be eliminated by informing all agents equally. For example, patients cannot be informed to the same extent as physicians in many aspects of medical treatments. Nor can medical insurance carriers possess the same knowledge as that possessed by patients about their health status, a situation that leads to adverse selection in the purchase of insurance. Given the inherent difficulty in correcting information asymmetry in health care markets, institutional designs are practically the only feasible mechanisms for mitigating its unwanted effects. For instance, establishment of a review committee on medical care procedures and costs would tend to discourage the overcharging of patients; equivalently, it would discourage provision of medical care of the quality not commensurate with the fee paid. Further, clear, enforceable regulations on what patients need to disclose to insurers about their health status in order for the medical insurance to be effective would reduce the phenomenon of adverse selection.
The problem of health care financing in low-income countries is best tackled by involving all the agents in the economy in its solution: the central government, the households, health care providers and the civil society. The problem is how to pay for health care in a manner that does not exclude anyone from its consumption. A game-theoretic perspective on this problem reveals that different agents make unique contributions to its solution. The central government, for example, coordinates a decentralized provision of health services, partly financed by direct monetary contributions by households (user fees) and partly by revenue from general taxation. The civil society associations play a crucial role in mobilizing funds which can be used to finance health care consumption by the poor; further, they facilitate non-cash financing of health care provided in non-market settings.

6.2 Conclusions and policy implications

The paper has several, far reaching conclusions. To begin, the conditions for efficient functioning of a system of user fees do not exist in many low-income countries, especially in rural areas. In particular, health insurance markets, which are crucial for effective functioning of curative health services markets are either missing or highly imperfect. In consequence, a system of fees, if applied as a principal mechanism for financing health services in these countries would lead to serious inefficiencies and inequalities in health care consumption. The consumption inefficiency arises from the fact that people who would be able and willing to hedge against risks of not receiving medical care would not do so because of nonexistence of insurance markets. The problem is that even though such people would always obtain medical care whenever it is needed, they would have to endure risks of not being able to afford medical treatment because of the absence of mechanisms for risk bearing. The consumption inequality is related to the absence of insurance markets: people who would otherwise prepay for medical care would be unable to afford treatment when it is needed.

In a low-income area, the inequality in health care would still persist even if health insurance markets were available because the poor would be unable to afford insurance premiums. Thus, government health services in poor countries should be financed primarily through the tax revenue. Nonetheless, a system of moderate user charges which have already been
implemented in many low countries should be retained. Further, the fees should be introduced where they do not exist. There are a number of reasons for this recommendation. First, the fees would encourage efficiency in service provision and use. Second, in some cases where travel and time costs of using the services are unimportant, the fees would discourage 'moral hazard behaviour' -- the tendency not to take precaution against illness because treatment is freely available in government health facilities. Third, since the tax-financed health care in government health facilities has the status of a commons, that is, treatment there is available free of charge to all, user charges would raise the cost of attending the facilities and thus discourage their overuse. As argued in the paper, tax-financed health care in government health facilities is a form of a public medical insurance scheme. Substantial additional public revenue can be raised to finance this scheme via presumptive taxation.

In implementing user charges, which are essentially a co-payment intended to rationalize use of the public medical insurance scheme, social safety nets should be designed to help people who would be unable to use the scheme because of the fees. As explained in the paper, in areas where fees are being introduced for the first time, health care demand can fall drastically even when the fees are quite moderate.

A second conclusion of the paper is that civil society associations facilitate non-cash financing of health care in non-market settings. This non-cash financing of health care, typically consists of the time contribution of the person providing the care. This care cannot be availed to the patient without the time donation by the provider. Thus, in contrast to a market setting, in a non-market context, health care provision is financed via a provider charge; the charge is analogous to a self-tax in non-monetary form. If expressed in monetary magnitudes, the sum of 'provider charges' in non-market health sector might be larger than the sum of 'user charges' in the market sector.

Furthermore, within the market health sector, membership in civil society associations can confer scale economies to individuals in the financing of health services; that is, group members can use group bargaining power to obtain services at reduced fees or to obtain medical insurance at preferential rates. There is need therefore to use fiscal instruments such as tax subsidies to encourage formation of civil society groups because they
help make market and non-market health services broadly available to the population.

The third, and a central conclusion of the paper is that strategic interaction among economic agents influences the outcome of a given method of health care financing. As an illustration, the outcome of a system of user fees, depends for instance, on how particular civil society entities such the trade unions or the mass media react to the fees *ex post*. If *ex post*, these groups oppose the implementation of fees -- because of conflict of interests between them and the central government, the effectiveness of the fees, e.g., in easing recurrent health expenditure could be considerably reduced. At worst, the fees could end up being withdrawn. Analytic or solution concepts from the theory of games can fruitfully be used to design acceptable user fee systems in situations of strategic interactions; needless to say, these are the commonplace situations in any sector of the economy. In a simple case where the central government and the civil society associations have equal bargaining power, the *rationalizable* conflict resolution is the one that averages the desires of the two parties. For example, the user fee implemented might be the average of the fees desired by each of the two parties. Indeed, Ellis and McGuire (1990) have demonstrated, in a seminal paper, how game-theoretic solution concepts can be used to resolve the conflict between patients and providers as to the optimal quantity of medical care to be supplied when the health services market is not clearing. In their simplified example, in which patients and providers are assumed to have equal bargaining power, the quantity supplied is a simple average of the quantities desired by the two parties. This result, and others obtained under the assumption of asymmetric bargaining power (Ellis and McGuire 1990) have a wide range of applications in the health policy arena of user fee designs and implementation.

The role of the central government in resolving conflict among fee implementing agencies such as the ministries of health and civil society associations, for instance, is indispensable; and so, a market-oriented reform need not imply a diminished role for the government. (Furthermore, in this analysis, the 'market', the 'government' or the 'civil society' are not alternatives in health care financing). The central government, as a 'super player' in the health care financing game may mitigate conflict among the actors concerned about the effects of the fees through several mechanisms. It may establish a forum for public discussion of fees, thereby allaying
unnecessary fears as to their effects; it may sponsor or undertake an information campaign concerning their necessity; or it may enact legislation providing support for persons unable to pay fees. These and related functions of the central government facilitate a productive play of the health care financing game by society.
REFERENCES


de Ferranti, David (1985), Paying for Health Services in Developing Countries. World Bank, Washington DC.


UNU/WIDER Working Papers

WP1 Food, Economics and Entitlements by Amartya Sen, February 1986
WP2 Decomposition of Normalization Axiom in the Measurement of Poverty: A Comment by Nanak Kakwani, March 1986
WP3 The Intertemporal Effects of International Transfers by Pertti Haaparanta, April 1986
WP4 Income Inequality, Welfare and Poverty in a Developing Economy with Applications to Sri Lanka by Nanak Kakwani, April 1986
WP5 Liberalization of Capital Movements and Trade: Real Appreciation, Employment and Welfare by Pertti Haaparanta and Juha Kähkönen, August 1986
WP6 Dual Exchange Markets and Intervention by Pertti Haaparanta, August 1986
WP7 Real and Relative Wage Rigidities - Wage Indexation in the Open Economy Staggered Contracts Model by Pertti Haaparanta, August 1986
WP8 On Measuring Undernutrition by Nanak Kakwani, December 1986
WP9 Is Sex Bias Significant? by Nanak Kakwani, December 1986
WP10 Adapting to Undernourishment: The Clinical Evidence and Its Implications by Partha Dasgupta and Debraj Ray, April 1987
WP12 The International Debt Problem - Prospects and Solutions by Stephany Griffith-Jones, June 1987
WP13 Walras' Law by Don Patinkin, June 1987
WP14 Technological Stagnation, Tenurial Laws and Adverse Selection by Kaushik Basu, June 1987
WP15 Undernutrition in Sub-Saharan Africa: A Critical Assessment of the Evidence by Peter Svedberg, June 1987
WP16 Controversies in Nutrition and Their Implications for the Economics of Food by S. R. Osmani, July 1987
WP17 Smallpox in Two Systems of Knowledge by Frédérique Apffel Marglin, July 1987
WP18 Gender and Cooperative Conflicts by Amartya Sen, July 1987
WP19 Africa and India: What Do We Have to Learn from Each Other? by Amartya Sen, August 1987
WP20 A Theory of Association: Social Status, Prices and Markets by Kaushik Basu, August 1987

WP21 A Theory of Surplus Labour by Kaushik Basu, August 1987

WP22 Some Reflections on Comparative Latin American Economic Performance and Policy by Albert Fishlow, August 1987

WP23 Post-Keynesian Theorists and the Theory of Economic Development by Sukhamoy Chakravarty, August 1987

WP24 Economic Reform in the USSR by Georgy Skorov, August 1987

WP25 Freedom of Choice: Concept and Content by Amartya Sen, August 1987

WP26 Ethiopian Famines 1973-1985: A Case-Study by Gopalakrishna Kumar, November 1987

WP27 Feeding China: The Experience since 1949 by Carl Riskin, November 1987

WP28 Market Responses to Anti-Hunger Policies: Effects on Wages, Prices and Employment by Martin Ravallion, November 1987

WP29 The Food Problems of Bangladesh by S. R. Osmani, November 1987

WP30 Internal Criticism and Indian Rationalist Traditions by Martha Nussbaum and Amartya Sen, December 1987


WP32 Non-Relative Virtues: An Aristotelian Approach by Martha Nussbaum, December 1987

WP33 Modernization and Its Discontents, a Perspective from the Sociology of Knowledge by Tariq Banuri, December 1987

WP34 Commodity Instability and Developing Countries: The Debate by Alfred Maizels, January 1988


WP38 Macropolicy in the Rise and Fall of the Golden Age by Gerald Epstein and Juliet Schor, February 1988
WP39 Profit Squeeze and Keynesian Theory by Stephen Marglin and Amit Bhaduri, April 1988

WP40 The Diversity of Unemployment Experience since 1973 by Bob Rowthorn and Andrew Glyn, April 1988

WP41 Economic Openness - Problems to the Century's End by Lance Taylor, April 1988


WP43 The Rise and Fall of the Golden Age by Andrew Glyn, Alan Hughes, Alan Lipietz and Ajit Singh, April 1988

WP44 The Food Crisis in Africa: A Comparative Structural Analysis by Jean-Philippe Platteau, April 1988

WP45 Famine Prevention in India by Jean Drèze, May 1988

WP46 A Model of Nutrition, Health and Economic Productivity by Peter Svedberg, September 1988

WP47 Undernutrition in Sub-Saharan Africa: Is There a Sex-Bias? by Peter Svedberg, September 1988


WP49 Social Security in South Asia by S. R. Osmani, December 1988

WP50 Food and the History of India - An 'Entitlement' Approach by S. R. Osmani, December 1988

WP51 Reform, Stabilization Policies, and Economic Adjustment in Poland by Grzegorz W. Kolodko, January 1989

WP52 Adjustment through Opening of Socialist Economies by Dariusz Rosati and Kalman Mizsei, January 1989

WP53 Reforming Process and Consolidation in the Soviet Economy by Andrei Vernikov, January 1989

WP54 Stabilisation and Reform in the Hungarian Economy of the late 1980s by Adam Török, March 1989

WP55 Economic System Reform in China by Zhang Yuyan, March 1989

WP56 Sectoral Balance: A Survey by Amitava Krishna Dutt, March 1989

WP57 Financial Markets and Governments by Robert Pringle, June 1989

WP58 Grassroots Strategies and Directed Development in Tanzania: The Case of the Fishing Sector by Marja-Liisa Swantz, August 1989
WP59 Defending the Right to Subsist: The State vs. the Urban Informal Economy in Tanzania by Aili Mari Tripp, August 1989

WP60 A Proposal for "COoperative Relief of Debt in Africa" (CORDA) by Jacques H. Drèze, Albert Kervyn de Lettenhove, Jean-Philippe Platteau and Paul Reding, August 1989

WP61 Limited Liability and the Existence of Share Tenancy by Kaushik Basu, August 1989

WP62 Black Markets, Openness, and Central Bank Autonomy by Tariq Banuri, August 1989

WP63 The Soft Option of the Reserve Currency Status by Amit Bhaduri, August 1989

WP64 Exchange Controls and Policy Autonomy - The Case of Australia 1983-88 by Andrew Glyn, August 1989

WP65 Capital Mobility and Policy Effectiveness in a Solvency Crisis. The Mexican Economy in the 1980s by Jaime Ros, August 1989

WP66 Quality of Life Measures in Health Care and Medical Ethics by Dan W. Brock, August 1989


WP68 Justice, Gender and International Boundaries by Onora O'Neill, August 1989

WP69 The Relativity of the Welfare Concept by Bernard M. S. van Praag, August 1989

WP70 Objectivity and the Science/Ethics Distinction by Hilary Putnam, August 1989

WP71 Distributing Health: The Allocation of Resources by an International Agency by John E. Roemer, August 1989

WP72 Explanation and Practical Reason by Charles Taylor, August 1989

WP73 International Capital Markets and the Limits of National Economic Policy by Gerald Epstein and Herbert Gintis, October 1989


WP76 Gap Disequilibria: Inflation, Investment, Saving and Foreign Exchange by Lance Taylor, October 1989
WP77 Soviet Economy: Opening up and Stabilization by Andrei Vernikov, October 1989

WP78 The International Debt Problem: Could Someone Please Explain It to Me? by Kaushik Basu, October 1989

WP79 Rural Women, Informal Sector and Household Economy in Tanzania by C. K. Omari, October 1989


WP81 Stabilization Policy in Poland. Challenges and Constraints by Grzegorz W. Kolodko, Marian Ostrowski and Dariusz Rosati, February 1990

WP82 Food Deprivation and Undernutrition in Rural Bangladesh by S. R. Osmani, February 1990

WP83 Modified Planned Economies at the Crossroads: The Case of Hungary by Kálman Mizsei and Adám Török, March 1990

WP84 The Environment as a Commodity by Partha Dasgupta, March 1990

WP85 Determinants of Female Labor Force Participation in the Middle East and North Africa by V. M. Moghadam, May 1990

WP86 Political Theories of Development Cooperation - A Study of Theories of International Cooperation by Lauri Siitonen, July 1990

WP87 Gender and Restructuring: Perestroika, the 1989 Revolutions, and Women by Valentine M. Moghadam, November 1990

WP88 Commodity Price Fluctuations and Macro-economic Adjustments in the Developed Countries by Walter C. Labys and Alfred Maizels, November 1990

WP89 Oil Driven Macroeconometric Model of Kuwait by Siddig A. Salih, William H. Branson and Yusuf H. Al Ebraheem, March 1991

WP90 Property Rights and Economic Reform in Uzbekistan by Simon Johnson and Bakhtior Islamov, September 1991


WP92 Accelerating Privatization in Eastern Europe: The Case of Poland by Jeffrey D. Sachs, September 1991

WP93 The Emergence and Stabilization of Extreme Inflationary Pressures in the Soviet Union by Ardo H. Hansson, September 1991

WP94 The Importance of Being Earnest: Early Stages of the West German Wirtschaftswunder by Ardo H. Hansson, September 1991
WP95 Understanding the "Informal Sector": A Survey by Madhura Swaminathan, December 1991


WP97 Managing Renewable Natural Capital in Africa by Siddig A. Salih, February 1992

WP98 Rural Poverty, Public Policy and Social Change: Some Findings from Surveys of Six Villages by Haris Gazdar, May 1992

WP99 Development and Patriarchy: The Middle East and North Africa in Economic and Demographic Transition by Valentine M. Moghadam, July 1992

WP100 Lush Fields and Parched Throats: The Political Economy of Groundwater in Gujarat by Bela Bhatia, August 1992

WP101 Utilities, Preferences and Substantive Goods by John C. Harsanyi, December 1992

WP102 The Ethiopian Famines, Entitlements and Governance by Derseh Endale, February 1993

WP103 External Imbalances, Famines and Entitlements: A Case Study by Derseh Endale, February 1993

WP104 Rural Markets, Food-Grain Prices and Famines: A Study on Selected Regions in Ethiopia by Derseh Endale, February 1993

WP105 Production Aspects of Russian Transition by Alexander Yu. Vorobyov, June 1993

WP106 Monetary Aspects of Russian Transition by Stanislav Zhukov, June 1993

WP107 The Entitlement Approach to Famine: An Assessment by S. R. Osmani, June 1993


WP110 Social Protection and Women Workers in Asia by Valentine M. Moghadam, June 1993

WP111 The Government Budget and the Economic Transformation of Poland by Alain de Crombrugghe and David Lipton, July 1993

WP112 Decentralized Socialism and Macroeconomic Stability: Lessons from China by Gang Fan and Wing Thye Woo, July 1993
WP113 Transforming an Economy while Building a Nation: The Case of Estonia by Ardo H. Hansson, July 1993

WP114 The J-curve is a Gamma-curve: Initial Welfare Consequences of Price Liberalization in Eastern Europe by Bryan W. Roberts, July 1993

WP115 The Restructuring Process of Rural Russian Karelia: A Case Study of Two Karelian Villages by Eira Varis, February 1994

WP116 Market Reforms and Women Workers in Vietnam: A Case Study of Hanoi and Ho Chi Minh City by Valentine M. Moghadam, July 1994

WP117 Sustainable Ecosystem in Africa: Managing Natural Forest in Sudan by Siddig A. Salih, December 1994

WP118 Employment-Based Safety Nets: Exploring an Alternative Approach to Limit the Adverse Consequences of Recurrent Droughts in Ethiopia by Derseh Endale, April 1995

WP119 The Economics of Complex Humanitarian Emergencies: Preliminary Approaches and Findings by E. Wayne Nafziger, September 1996

WP120 Health Effects of Market-Based Reforms in Developing Countries by Germano Mwabu, September 1996


WP123 Promoting Education within the Context of a Neo-patrimonial State: The Case of Nigeria by Daniel Edevbaro, January 1997


WP125 Privatization, Asset Distribution and Equity in Transitional Economies by Juha Honkkila, February 1997

WP126 Economic Shocks, Impoverishment and Poverty-Related Mortality during the Eastern European Transition by Renato Paniccià, March 1997

WP127 User Charges for Health Care: A Review of the Underlying Theory and Assumptions by Germano Mwabu, March 1997

Please note that certain WP issues are no longer available. For further information, please contact UNU/WIDER Publications at the address given at the beginning of this publication.