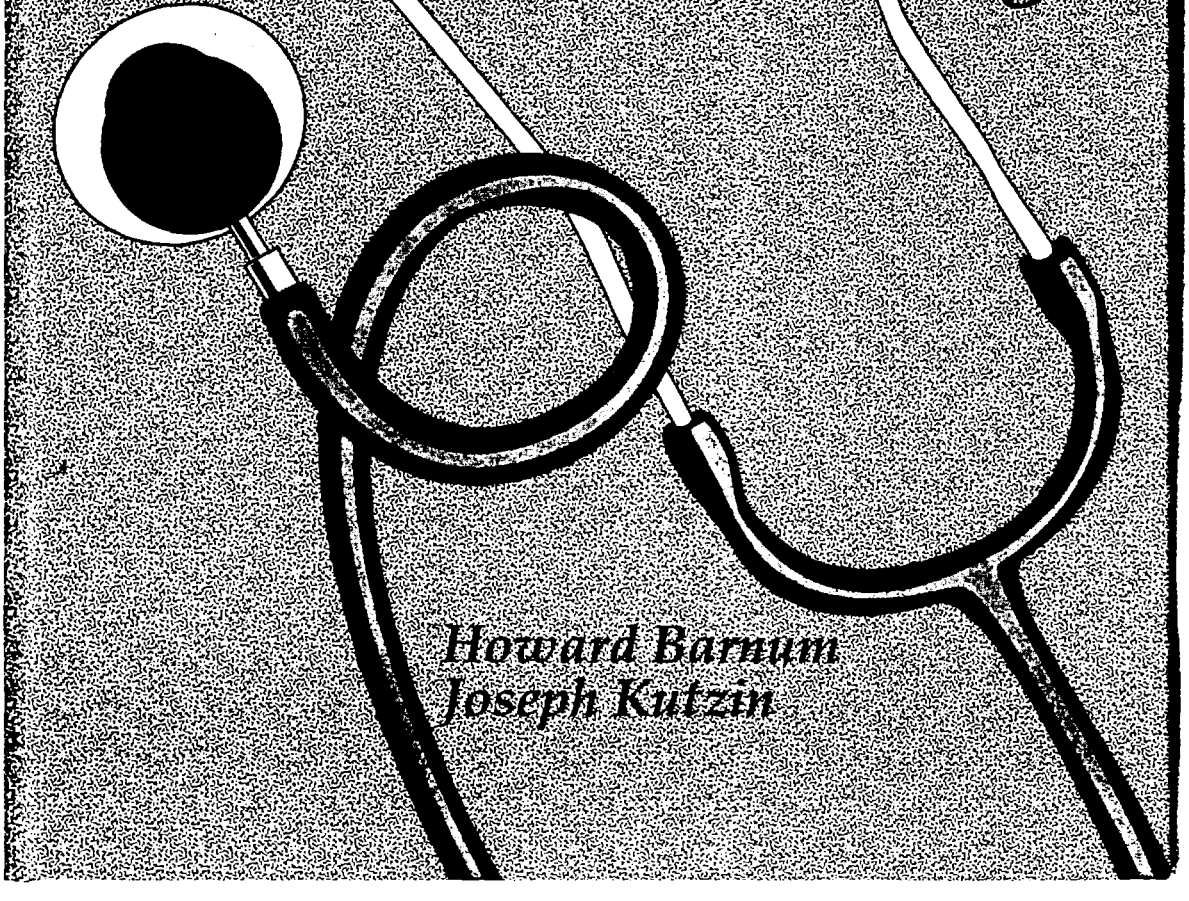
Public Disclosure Authorized

Public Hospitals in Developing Countries

Resource Use, Cost, Financing



Howard Barnum
Joseph Kutzin

Public Hospitals in Developing Countries

A World Bank Book

Public Hospitals in Developing Countries

Resource Use, Cost, Financing

Howard Barnum
and
Joseph Kutzin

Published for the World Bank
The Johns Hopkins University Press
Baltimore and London

Manufactured in the United States of America
First printing April 1993

The Johns Hopkins University Press
Baltimore, Maryland 21211, U.S.A.

The complete backlist of publications from the World Bank is shown in the annual *Index of Publications*, which contains an alphabetical title list and indexes of subjects, authors, and countries and regions. The latest edition is available free of charge from Distribution Unit, Office of the Publisher, The World Bank, 1818 H Street, N.W., Washington, D.C. 20433, U.S.A., or from Publications, The World Bank, 66 avenue d'Iéna, 75116 Paris, France.

Library of Congress Cataloging-in-Publication Data

Barnum, Howard.
Public hospitals in developing countries : resource use, cost, financing / Howard Barnum and Joseph Kutzin.
p. cm.
"Published for the World Bank."
Includes bibliographical references and index.
ISBN 0-8018-4532-7
1. Hospitals—Developing countries—Finance. 2. Hospitals—Developing countries—Cost control. I. Kutzin, Joseph. II. International Bank for Reconstruction and Development. III. Title.
[DNLM: 1. Developing Countries. 2. Financial Management, Hospital. 3. Hospitals, Public—economics. 4. Hospitals, Public—utilization. WX 157 B263p]
RA971.3.B37 1993
338.4'336211'091724—dc20
DNLM/DLC 92-49341
for Library of Congress CIP

Contents

Foreword

Hospitals consume the largest share of government health resources, yet until recently, they have not been a focus of health policy research in developing countries. The attention paid to community-based and other nonhospital primary care interventions was appropriate, of course, because of the cost-effectiveness of these measures compared with those delivered in hospitals. But this should not have led to analytical neglect of hospitals. It is precisely because hospitals account for such a large share of health expenditure that improvements in their efficiency can yield tremendous benefits for the entire sector. Moreover, strong hospitals at the first referral level are vital inputs to the success of an integrated primary health care strategy. For example, the strengthening of surgical and special obstetric services at the referral hospital is a key component of the strategy for reducing maternal deaths in developing countries.

Most of the studies that are discussed in this book date back only to 1987, and data and analyses presented herein will prove useful to both researchers and policymakers. The cases reviewed in this book illustrate analyses that hospital managers in developing countries should carry out in their own institutions. In addition, the authors highlight a number of unanswered questions that can form the basis for further studies, either national or cross-national.

By improving the efficiency of hospitals in developing countries, a greater number of persons can be served with available resources. By ensuring that hospitals provide appropriate support to the primary care level and that they are well integrated into national health systems, their effectiveness can be much enhanced. This is of vital importance in this era of worldwide fiscal constraints.

I commend this book to all engaged in improving the health of people living in developing countries.

Adetokunbo O. Lucas
Director, International Health Programs
Department of Population and International Health
Harvard School of Public Health

Acknowledgments

This book owes much to the interaction, help, and encouragement of many colleagues during a period of years. At an early stage in our formulation of the research effort, when hospitals were a comparatively neglected topic of research in the health sectors of developing countries, the support and interest of Nancy Birdsall, Anthony Measham, and Dean Jamison were invaluable. In order to identify the issues and establish a manageable focus, we began our research with a set of background papers prepared in 1986 and 1987 by Kenneth Chomitz (1986), David Dunlop (1986), Julio Frenk, Enrique Ruelas, and Avedis Donabedian (1989), Andrew Green (1987), Dominique Jolly (1987), Anne Mills (1990a and 1990b), and Godfrey Walker (1986). As indicated in the References, some of these studies have since been published and are generally available. Our examination of these broad overviews was followed by studies in specific countries: China (Barnum 1989), Ethiopia (Bitran-Dicowsky and Dunlop 1989), Indonesia (Djuhari and others 1988; Barnum 1987b), Malawi (Mills 1991), and the Dominican Republic, Honduras, and Jamaica (Lewis 1990). The section on statistical cost functions in chapter 3 is based partly on ongoing research we are conducting with Adam Wagstaff on the characteristics of hospital cost functions.

In addition to the activities commissioned specifically for this book, the findings and recommendations are based on a detailed survey of the relevant literature, both published and unpublished. Sector work conducted under World Bank auspices has contributed greatly to our research. Although we have drawn information from many World Bank reports, the work managed by John Briscoe and William McGreevey in Brazil, Richard Bumgarner in China, Nicholas Prescott in Indonesia, and Hazel Denton in Nigeria deserves specific mention. Many non-Bank studies were also used. Of particular note is the work carried out under two activities funded by the United States Agency for International Development: the Health Care Financing in Latin America and the Caribbean (HCF/LAC) project, managed by Dieter Zschock; and the Resources for Child Health (REACH) Project, managed by Gerald Rosenthal.

In May 1989, a workshop entitled "Hospital Resource Use and Cost Containment" was held at the World Bank. Presentations on hospital costs, financing, and health service alternatives were made by Ricardo Bitran-Dicowsky, David Dunlop, Charles Griffin, Maureen Lewis, Anne Mills, Germano Mwabu, Julia Walsh, Annemarie Wouters, and Mary Young. José-Luis Bobadilla, Richard Bumgarner, Philip Musgrove, Mead Over, and Donald Shepard added greatly to the discussion of these issues. Following the workshop, we began work on this book.

Many people reviewed parts or all of earlier drafts of the manuscript, some on several occasions. We are most grateful for their written comments as well as contributions made during meetings held to discuss the first draft in August 1990 at the World Bank. Special acknowledgment is due to Martha Ainsworth, Ricardo Bitran-Dicowsky, José-Luis Bobadilla, John Briscoe, Richard Bumgarner, Guy Carrin, Andrew Creese, Carlos Cruz-Rivero, Hazel Denton, David Dunlop, Guy Ellena, Julio Frenk, Willy de Geyndt, Fred Golladay, Charles Griffin, Davidson Gwatkin, Salim Habayeb, Jeffrey Hammer, Jean-Louis Lamboray, Maureen Lewis, Patricio Marquez, Jo Martins, Anthony Measham, William Newbrander, Mead Over, Donald Shepard, Anne Tinker, Louis Vassiliou, Adam Wagstaff, Julia Walsh, Marcia Weaver, Vivian Wong, and Annemarie Wouters. The exacting readings given the manuscript by four anonymous referees commissioned by the World Bank's Office of the Publisher were also invaluable and resulted in important revisions. Any remaining omissions and errors are our responsibility.

We are grateful to Joanne S. Ainsworth for a splendid job of editing this book and to Leonila Jose, who provided excellent secretarial assistance throughout all stages of preparation of the manuscript.

1. Introduction

This book examines economic and financial issues of hospital resource allocation with the objective of contributing to policies that will improve the use of public sector funds by hospitals. More specifically, the book concerns (a) the allocation of health sector resources between hospitals and nonhospital alternatives, (b) the internal efficiency of hospital operations, and (c) effective and equitable cost-recovery policies for hospitals.

Hospital operating expenses are at the core of the growing gap between required and available resources in the health sector of many countries. Hospitals receive the lion's share of public recurrent resources in health. Although the actual percentage varies from country to country, it is common for 50 to 80 percent of public sector health resources, in money and trained personnel, to be used in hospitals. The remaining resources are used for preventive care, infectious and parasitic disease control programs, nonhospital maternal and child health programs, and other services that have been identified as generally more cost-effective than hospital care in countries with limited health resources. Reviews of the health sector in many countries suggest that these large recurrent expenditures on hospitals involve a great waste of resources because of the inappropriate allocation of funds within the health sector and the technical and managerial inefficiency within hospitals.

Notwithstanding the large share of government health resources allocated to hospitals, much of health research has focused on maternal and child health programs and first-line services at the periphery rather than on hospitals and higher-level curative care. The emphasis on nonhospital interventions reflects the limited results produced by capital-intensive hospital projects. The avoidance of capital expenditure on hospitals in health sector projects has been, and generally continues to be, desirable, but the neglect of hospitals in sector research has been unfortunate, because research can be influential in changing the role of hospitals and the flow of resources in the sector. In addition, the emphasis on nonhospital interventions may have overshadowed the important potential contribution of hospitals to the integral health system.

Just five or six years ago there was relatively little information on hospital economics in developing countries. Recent studies, however, sponsored by ministries of health in collaboration with such donors as the World Health Organization (WHO), the United States Agency for International Development (USAID), and the World Bank, have begun to fill the void in hospital research. Although our knowledge of hospital costs, finance, and resource use remains incomplete, the new research provides important data that are needed to develop improved hospital policy. This book draws widely on this research to support the analysis presented.

The Scope of the Book

The overarching viewpoint of this book is economic. Thus we examine the main questions of hospital efficiency within the framework of the economics of production and costs. Full answers to these questions must also involve the fields of hospital management, organization, and personnel planning, but to keep the book manageable we have restricted the scope to economic and financial issues. Economic analysis can help to identify the areas in which changes are needed even though implementation of the changes ultimately requires noneconomic analysis. Also, some important policy areas, especially those related to incentives and pricing, directly involve economics and finance.

The economic view of hospitals provides an important perspective, but it is not all-encompassing. This perspective is subsumed by a broader public health view of hospitals, which considers hospitals in relation to the epidemiology of the population and the structure of other health services. Public health policy will reflect the way in which health authorities attempt to meet health needs, as determined by a population's epidemiology and demographic structure. The level of preventive activities, the emphasis on basic care provided in health centers and through outreach programs, referral policies, and the choices made by consumers of health services all act to determine the case mix faced by a given type of hospital. Treatment norms within the hospital (and in the medical community in general), administrative policies, and staff composition and incentives determine the procedures used for cases of a given type. Although we discuss many aspects of public health policy in our economic approach to hospitals, we do not attempt an in-depth analysis of all aspects of the health system. The policy recommendations reached from the economic analysis of hospitals, however, do have implications for the entire health sector.

The scope of this book is limited primarily to public hospitals and to public programs that finance hospital care, irrespective of whether such

care is delivered by governmental or private providers. This limitation is consistent with the objective of the book, which is to contribute to public sector efficiency. It also has kept the book within manageable bounds by restricting the information needed to the relatively more accessible public hospital data. Information on private sector hospitals is difficult to obtain because very few countries have any system of regular reporting to a central authority on the details of private sector hospital operations. Nevertheless, nonprofit hospitals operated by nongovernmental organizations (NGOs) as well as profit-seeking hospitals play an important role in the provision of services in most developing countries. Public and private sector hospitals inevitably interact in the market for services, and this interaction is taken into account in the policy discussion.

The Issues

Correcting the problems with hospital resource use does not simply mean allocating a greater proportion of development funds at the margin to nonhospital projects. In order to control hospital expenditures and improve the efficiency, management, and role of hospitals in the health sector, we need a better understanding of the issues underlying the allocation of health sector resources to hospitals. Research on questions of hospital economics and finance is needed to gain this understanding and help formulate strategies for more effective hospital resource use. Some of the questions that must be answered follow:

- *What share of financial and personnel resources do hospitals absorb? Who uses hospitals? Given an appropriate criterion, such as maximum health benefit, is the share of public health sector resources absorbed by hospitals optimum?*

  The actual share of resources going to hospitals is not clearly known in many countries, but aggregate government data, together with special detailed studies in some cases, do suggest substantial differences among countries in the allocation of public sector health resources. The amount of health sector resources absorbed by hospitals in contrast to that taken in by nonhospital programs can have implications for the success of primary health care and the overall effectiveness of the public health sector. The provision of hospital services also has implications for the distribution of benefits from public resources.

- *What are the levels of hospital average and marginal costs in selected regions and countries? How can these unit costs be measured? What are the relations among unit costs, occupancy rates, length of stay, and quality of*

services? What are the determinants of hospital unit costs? To what extent can hospital unit costs be reduced and efficiency increased through adjustments to factor mix and changes in length of stay? How does efficiency vary with type and size of hospital?

Knowledge of unit costs is needed to assist planning for recurrent budgets, as an indicator of efficiency, and to inform pricing of services. Costs are not known with any accuracy for most hospitals in developing countries. Recent studies have demonstrated the feasibility, however, of reconstructing hospital expenditures to obtain estimates of unit costs. This information can be used to suggest means of increasing hospital efficiency through adjustments to factor mix, changes in length of stay, and improved use of the referral system.

- *Is cost recovery feasible and desirable in hospitals? What principles should determine the level of hospital fees? Can practical guidelines be developed? Should risk-sharing mechanisms be used to finance hospitals? What institutional features should be included in insurance programs?*

Decreased availability of government recurrent revenues has put new pressure on government health officials to look for alternative sources of funds. Cost recovery has been promoted by donors and by government ministries but with little practical guidance on how to set fees and for what services. Cost recovery in health could start most readily with the use of fees for hospital services, where, because of administrative capacity and concentration on curative care, the use of fees is more practical than it would be for lower-level services. The appropriate level of these fees, however, is open for discussion.

- *Can the functioning of the referral system be improved in order to increase access to care and the economic efficiency of the health sector? Are there practical, lower-cost alternatives to the delivery of health care now delivered through secondary and tertiary hospitals?*

The high cost of technically complex hospital services has motivated a search for less complex alternatives that could provide acceptable care at lower cost. Some of the possibilities considered have included redefining the role of district-level hospitals, use of home care, and greater reliance on short-term stays and outpatient care.

Organization of the Book and Overview of the Findings

The organization of the book flows directly from the issues set out above. The magnitudes of hospital use of public sector health resources and hospital unit costs are identified, and then possible financing and health sector service alternatives are considered. A brief overview of the book is given below. More detailed summaries are provided at the end of chapters 2–6. Chapter 7 recapitulates our principal findings and

recommendations and concludes with suggested topics for future research.

Hospital Resource Use

The magnitude and patterns of hospital resource use are examined in chapter 2. As expected, we found that in almost all countries hospitals receive the largest share of government health resources and that, within the hospital subsector of many countries, tertiary hospitals receive a large share of both financial and skilled personnel resources in relation to that received by district hospitals. Furthermore, we found that hospital services are used disproportionately by urban populations, and there is evidence (although less clear) to suggest that hospital services are used less by the poor. Another finding is that not only do adults and the elderly use hospitals more than their share of the population would indicate, but they are also more expensive to treat. Thus as the population ages, pressure on hospital resources will increase if services continue to be delivered in the current manner. Finally, the chapter includes a survey of available studies on the cost-effectiveness of individual health interventions. These studies confirm that, in low-income economies, nonhospital interventions are more efficient in dealing with the prevalent health conditions. Although not surprising, these findings are important because they emphasize the continuing potential problem of extensive use of health sector resources in hospitals.

In addition to these more evident results are some findings that lead to less evident conclusions. One surprising result is the lack of a clear inverse correlation between the hospital share of resources and success in achieving primary health care objectives. In countries that have achieved important gains in reducing infant mortality during the last twenty years, such as Sri Lanka and China, both a high absolute level of support for the health sector and the involvement of district-level hospitals in a broad sectoral strategy have contributed to the success of primary health care programs even though hospitals absorb more than 60 percent of public recurrent health spending.

More important than the share of public health resources absorbed by hospitals is the use of the resources within the hospital subsector and the overall composition of health programs. In countries with successful primary health care programs, there is an emphasis on district-level hospitals rather than large tertiary facilities. Additionally, as the level of income increases, the relative importance of hospital services increases. Thus, the optimum proportion of health sector resources absorbed by hospitals must be determined by assessment of the epidemiological needs and the available personnel skills, the financial resources, and the structure of the hospital subsector of individual countries.

Hospital Costs

Hospital unit costs and their implications for policy are discussed in chapter 3. This section relies both on accounting-based studies and on statistical estimation as complementary tools to examine internal efficiency issues and to generate a cost basis for planning and cost-recovery policies. The book does not give full answers to all the questions raised above. Information is provided on the magnitude of costs, the methodologies for measuring costs, and the relationships between the various service indicators. But answers to the questions of the relation of hospital cost to efficiency remain tentative, primarily because of the difficulty in measuring the quality of hospital services. Quality issues have not received satisfactory treatment in the literature on hospital cost and are an important topic for future research. With the caveat concerning the issue of quality, we can summarize some of the chapter's findings.

An important fact demonstrated by the survey of accounting studies of average costs is that such studies are feasible in low-income countries. Collection and analysis of the data required to calculate average costs can be made a routine hospital activity with the objective of improved planning, management, and budgeting. The estimated average costs demonstrate that tertiary-level hospitals have substantially higher costs per patient than do more basic hospitals. This situation is to be expected, but the treatment by tertiary facilities of cases not requiring specialized care reflects a waste and misuse of resources. The difficulty, of course, is in providing credible services at the district level that can support a more rational use of referral facilities. In chapter 3 (and chapter 6 as well) we discuss the importance of improving the quality of district facilities.

There have been very few statistical studies of developing-country hospital costs, and it is thus premature to generalize from their findings. The cost functions discussed in chapter 3 do not show evidence of long-run economies of scale or of economies of scope.[1] These findings are consistent with similar studies of hospital costs in industrial countries. Clearly, however, more examples from developing countries are needed to improve our understanding of the hospital production process in this context.

Many relations between service statistics and efficiency have been identified. A cross-country survey reveals that occupancy rates, bed turnover rates, and average lengths of stay vary considerably across countries and among levels of hospitals.[2] Low turnover and occupancy rates, especially among lower-level facilities, are often attributable to the poor quality of services caused by insufficient drugs and other supplies and a lack of skilled staff. Analysis suggests that, with plausible demand response, higher occupancy and possibly lowered average cost can be compatible with greater expenditures on facilities to improve quality.

Another finding is that low turnover and long average lengths of stay contribute to high costs per case treated. Much of this problem is managerial and requires careful studies if solutions are to be proposed in individual cases. Some potential causes are poor scheduling of diagnostic services or surgery, outmoded treatment protocols, and misuse of secondary and tertiary facilities for extended care or convalescence. The discussion in chapter 3, as well as in chapters 4 and 5, also suggests that poor hospital performance, as revealed by service statistics, can relate to adverse incentives provided by the structure of hospital financing.

Financing

Hospital cost recovery is considered in chapters 4 and 5. In these chapters we build on the unit cost discussion and examine efficiency, equity, and revenue implications of the use of fees and insurance for hospital services. Chapter 4 provides a general framework for the analysis; here we review the deficiencies in the private market that lead to public intervention and set out the efficiency, equity, and revenue objectives of financing policy. In chapter 5 we review current cost-recovery policies in a variety of country settings.

Chapter 4 contains some principles for setting hospital prices: (a) fees should be consistent with ability to pay, (b) fees should provide signals that promote economic efficiency, (c) fees and the quality of services should be linked, (d) fees should be subsidized for services that have externalities, are public goods, have associated informational deficiencies, or are merit goods.[3] A full answer to the question of how to set fees requires information on both sides of the market—household demand and the hospital cost of supplying services. Studies of hospital unit costs provide an important supply-side component to help determine appropriate fees, whereas studies of service use response to income and price levels provide required demand-side information. Application of recent public enterprise economics integrates demand- and supply-side information to allow the development of guidelines for fee determination that make explicit the roles of costs and demand response. Briefly, prices should reflect price and income elasticities as well as costs. Services that are more income elastic (that is, a given percentage increase in income is accompanied by an even greater percentage change in the quantity of service demanded) should have a greater proportion of costs reflected in their price than services that are income inelastic. The benefits of applying this rule are improved equity and efficiency.

The application of pricing rules must be tempered according to the institutional setting in which financing takes place. The use of fees, risk-sharing arrangements, management and organization, and the overall structure of the health system all greatly affect client and pro-

ducer incentives and performance in individual countries. In chapter 5 we review the systems used to finance health care in several countries and set forth some practical additions to the pricing principles outlined above. These additions can be abbreviated here: timely fee adjustments should be an integral part of the system, fee structures should be simple, and exemptions should be limited to those granted on equity grounds or for services that provide benefits well beyond the patient. The chapter also includes some suggested guidelines for risk sharing: cost containment requires active involvement by the insuring institution to encourage providers to behave in a cost-effective manner; cost sharing is needed in all risk-sharing schemes to provide consumer incentives for efficiency (but these appear to be less effective than provider incentives); prepaid capitation schemes (for example, health maintenance organizations) are practical in circumscribed communities. Publicly sponsored provision of insured hospital services for only a small proportion of the population should be done with great caution to avoid inequitable distributional consequences.

Health Sector Alternatives

Lower-cost alternatives to hospitals are considered in chapter 6. In this chapter we examine the role of hospitals in the wider network of health service providers and suggest that there is ample scope for improving the overall cost-effectiveness of the sector through a reallocation of resources in favor of relatively low-cost providers. The suggested approach is to begin with an epidemiological picture of the population, delineate several viable packages of interventions for meeting health needs, and cost the alternative packages. The alternative that would yield the maximum health benefit for the available resources should be chosen. It is expected that a choice made in this manner would result in a network of providers and an allocation of resources that would encourage treatment of frequently occurring, low-cost conditions in the least expensive, yet capable, setting. Alternative treatment settings would include care in the home and community and noncomplex facilities such as health centers, which are often considered to be important components of a health referral system. Innovative alternatives not typically thought of in this manner, such as outpatient surgery facilities, could also be included if they are found to be cost-effective. As a result of the use of these alternatives, the case mix of hospitals of varying complexity would change so that, for example, tertiary hospitals would primarily treat patients requiring highly specialized care rather than using substantial resources for patients who could be cared for in less expensive facilities.

Practical ways in which the development of alternatives to hospital care can be supported are discussed in the chapter. It is most important to improve quality at lower-level facilities so that the population will not bypass them in favor of hospitals. To do this would require reallocating resources in favor of noncomplex facilities, perhaps away from hospitals. Another suggested policy is to coordinate the management of all providers at the district level, including the district hospital. The district hospital can also support primary care provided in lower-cost settings with training programs, assistance with logistics, and provision of credible diagnostic backup. Finally, we discuss alternatives to expensive hospital inpatient care that may be useful in many developing countries. These include treatments that can substitute for hospitalization and the use of lower-level facilities for treatment and palliation of chronic diseases.

Notes

1. See chapter 3 for definitions of these and other characteristics that can be derived from statistical cost functions.

2. See chapter 3 for definitions of these measures of service performance.

3. These terms are all aspects of market failure. A discussion of market failure and definitions of these terms are given in chapter 4.

2. *Patterns of Hospital Resource Use*

A change in health policy can affect the balance between services provided by hospitals and those provided by nonhospital facilities, and it can alter the share of government health sector resources used by hospitals. Interest in the share of health sector resources used by hospitals is motivated by the need to determine whether the allocation of resources within the health sector is economically efficient. An allocation of health sector resources among alternative activities is economically efficient if there is no possible reallocation that will improve health status. Unfortunately, appropriate measures of health status are not easily defined for the health sector in general and are especially hard to define for hospitals with their complex multiple outputs. Although hospitals are an essential part of any health system, the optimum allocation of resources is not obvious but is open to debate. Considerable analysis and discussion during the last twenty years have led to a rough consensus about the cost-effectiveness of primary curative care and preventive services compared with that of hospital services. Health planners generally agree that in countries with low levels of per capita gross national product (GNP) and high rates of mortality, the most effective allocation of resources would result in a lower share of public resources committed to hospitals than in countries with high GNP and low mortality.

Accordingly, the share of government health resources going to hospitals is a rough indicator of the structure and emphasis within the health sector. A relatively low share suggests an emphasis on primary health care and concern with reaching rural populations, and a high share suggests an emphasis on curative care and concern with urban health. This interpretation of the hospital share, however, should be applied cautiously to any specific country because the types of facilities defined as hospitals vary across countries and, apart from definitional issues, the absolute level of resources flowing to the health sector and the quality

of health services are also determinants of the structure and effectiveness of the health sector.

Despite the consensus that a greater share of public sector health resources should be allocated to nonhospital activities than is currently the case in many countries, the level of the appropriate percentage is not obvious. The provision of better quality and increased quantity of primary health care services need not conflict directly with the existing allocation of resources. Primary curative care is one type of service currently provided in many hospitals, and as discussed in chapter 6, hospitals can also provide a variety of preventive, palliative, and rehabilitative services and be structured so that they integrate with and provide support for nonhospital services. Hospitals are relatively high-cost institutions, especially with regard to investment, but primary health care outreach activities, although relatively low cost in terms of investment, can entail large recurrent costs. What is important is to obtain a balance of services throughout the health sector, with each investment being examined on the margin for its cost-effectiveness and contribution to sectoral efficiency.

Interest in the share of public sector health resources going to hospitals is also motivated by a concern with equity. There is no universal definition of equity because it is a normative concept that is defined in the context of prevailing social and moral values. But again based on a consensus view, equity might be defined as equal access across population and income subgroups to health services covering basic health needs. Using this minimum definition, equity could be enhanced with a progressive distribution of the burden of payment across income groups. Just what is included in basic health needs is only roughly defined, but at a minimum it encompasses safe childbirth, prevention and treatment of the main child health problems, protection from common tropical diseases, and access to first aid for accidents. Hospitals are not the most effective way to deal with many of these needs. In addition, hospital facilities are commonly distributed unequally across geographic areas and used unequally by different income groups. Thus, a relatively inequitable allocation of health sector resources is suggested by a very high percentage of recurrent health expenditures being absorbed by hospitals.

This chapter is largely descriptive of observed patterns of government resource allocation in the health and hospital sectors. The data are too variable in definition across countries to support strong policy prescriptions. Also, policies to reallocate sectoral resources in a particular country should be supported by detailed studies addressing the specific institutional issues of the setting. For these reasons we confine ourselves to identification of broad characteristics of hospital resource use and the potential policy implications of these patterns.

In the first two sections of this chapter we review cross-country information on the use of resources by public hospitals, first, with regard to the distribution of resources within the health sector and, second, with regard to the distribution of resources within hospitals. In the third section we assess data on the use of hospital services across income groups, geographic location, age, and disease types and consider the implications of existing use patterns for equity and future resource costs. In the fourth section of the chapter we review information related to the consensus view on the cost-effectiveness of hospital services. The chapter thus describes the resources used by hospitals and then turns to a normative evaluation of hospital resource use. The final section summarizes the principal findings of the chapter.

Distribution of Resources within the Government Health Sector

Hospitals compete with other health facilities and programs for both recurrent and capital resources. The amount allocated to hospitals must be further redistributed among different levels and types of hospitals. These allocational issues are described below.

Recurrent Resources

In most developing countries the distribution of recurrent resources within the government health sector strongly favors hospitals. Quantitative comparisons across countries are difficult because of variation in the definition of hospitals and a lack of comprehensive information available from the private sector and from different levels of government within the public sector. Broadly, the available data do reveal that a high percentage of government health resources are assigned to hospitals. Among the twenty-nine developing countries listed in table 2-1, over half spend more than 60 percent of recurrent health budgets on hospitals, and about two-thirds spend 50 percent or more. Only four countries distribute less than 40 percent of public sector health resources to hospitals. However, the implications of the hospital share for the success of primary health care (PHC) programs and for national health indicators are not clear and must be analyzed at the country level. Nepal, which has the lowest hospital share of any country in table 2-1, has a low level of absolute support for health services and notable problems in implementing quality primary care programs (World Bank 1989). It also has a low level of health status, reflected in an infant mortality rate (IMR) of more than 125 per 1,000 live births in 1987. Indonesia, also a low-income country, has been more successful at implementing primary care programs (Prescott 1991) and has a higher overall level of support for

Table 2-1. Share of Hospitals in Total Public Recurrent Health Expenditure
(percent)

Country	*Hospital share*	*Year*	*Source*
Bangladesh[a]	61	1987	Griffin 1989
Botswana	49	1984	
Brazil	68	1984	
Burundi	66	1986	
China[b]	61	1987	
Colombia	67	1984	
Côte d'Ivoire[c]	46	1984	Vogel 1988
El Salvador[d]	62	1985	Fiedler 1987
Ethiopia	49	1983–84	
The Gambia	45	1985–86	
Indonesia	37	1985–86	
Jamaica[e]	72	1986–87	Kutzin 1989
Jordan[f]	75	1987	
Kenya	73	1985–86	
Lesotho	74	1986–87	
Malawi[g]	81	1985–86	
Mexico[h]	58	1986	
Mozambique[i]	36	1987	
Nepal[j]	25	1987–88	
Niger	30	1988	Budd 1989
Papua New Guinea	45	1986	Newbrander 1987
Philippines[f]	71	1985	
Senegal	50	1982	Vogel 1988
Somalia	70	1989	
Sri Lanka	70	1986	ADB 1987
Swaziland	52	1983–84	
Turkey[k]	63	1987	
Uganda[l]	43	1982–83	
Zimbabwe[m]	54	1987	
OECD mean[n]	54	1980s	OECD 1987

a. Includes Upazila Health Complexes (primary-level hospitals).

b. Includes all expenditures by central and provincial governments on hospitals, hospitals of traditional Chinese medicine, and township hospitals.

c. Percentage reflects mean of 40–51 percent range reported by Vogel.

d. Percentage may reflect hospitals as a share of all facility-related health expenditures, rather than of total public sector health expenditure.

e. Includes mental health hospital and the share of drugs and other medical supplies allocated to hospitals from central administration.

f. Excludes health expenditures by public sector other than Ministry of Health.

g. Percentage reflects share of expenditure devoted to "curative" services, which overstates the hospital share to some extent.

h. Includes capital and recurrent public health expenditures.

i. Recurrent budget estimates.

j. Total (capital and recurrent) budget devoted to secondary and tertiary care and to hospitals with fewer than fifty beds; excludes amounts budgeted for military and mission hospitals.

k. Includes expenditures by the social insurance organization.

l. MOH expenditures (uncertain if capital included) with medical stores proportionally reallocated to facilities (may understate hospital share); some health center services included.

m. Includes health expenditures by MOH and municipal or local government health authorities; excludes central government health expenditures by other ministries.

n. Institutional health care, includes nursing homes, does not include ambulatory care.

Source: World Bank sector reviews and appraisal reports, except as noted above.

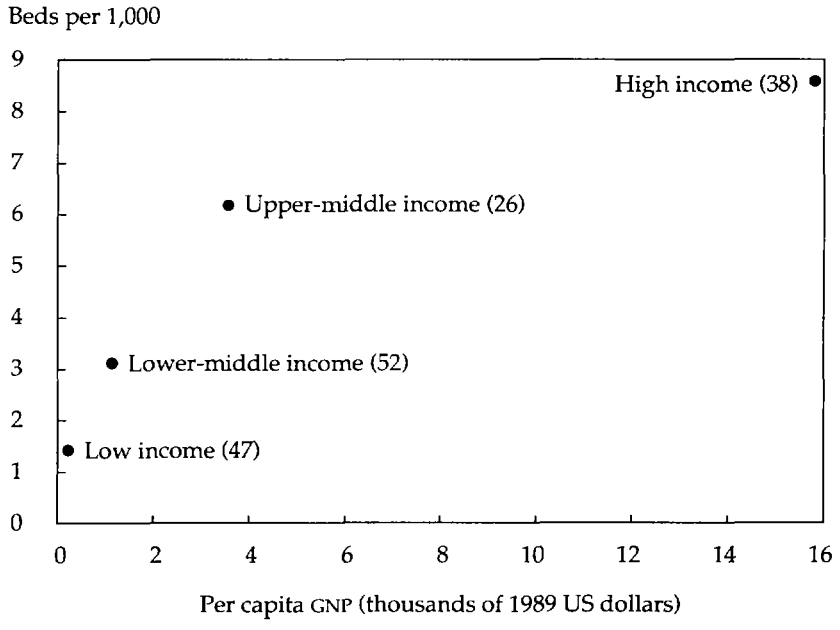
health services. Indonesia's IMR was 70 in 1987. At the other extreme, eight countries in table 2-1 devote 70 percent or more of public health resources to hospitals; among them Lesotho, Malawi, and Somalia have IMRs of more than 100, and Jamaica, Jordan, the Philippines, and Sri Lanka have IMRs below 50 (see the appendix table at the end of the chapter for IMR data).

The use of beds per capita as a measure of the intensity of hospital resource use is common but is fraught with difficulties. The definition of a hospital bed varies across countries. Most developing countries do not distinguish between acute beds (used for short-term stays of less than, say, thirty days) and long-term beds, although some countries do distinguish long-term specialty hospital beds, such as those for tuberculosis, leprosy, and psychiatric care. Even more troublesome are occasional large differences between the number of beds officially installed and the number of operational beds. A breakdown between acute and chronic beds, and confirmation that the beds given are operational, would add greatly to the interpretation of the comparative bed ratios, but this information is not available for many countries. Thus, bed ratios give only a gross measure of the availability of hospital services.

Generally, there is a strong relation between the availability of hospital beds per capita and the level of per capita GNP (figure 2-1). The number of beds per capita increases by about 4 percent for every 10 percent increase in per capita income.[1] The structure of the health sector in terms of the balance of resources between hospitals and other health service alternatives, however, is not as easily explained as the overall availability of hospital beds.

Despite the consensus that poorer countries should devote a larger share of public sector health resources to nonhospital interventions, available evidence suggests that there is no clear relation between the hospital resource share and the level of per capita national income. This lack of relationship is clear in the scatter of points relating the hospital share of government health expenditure on the Y axis and GNP per capita on the X axis in figure 2-2, which indicates that there are a number of poor countries with relatively high hospital shares. The high hospital shares of some poor countries may be explained (in part) by the nature

Figure 2-1. Hospital Beds per 1,000 Population

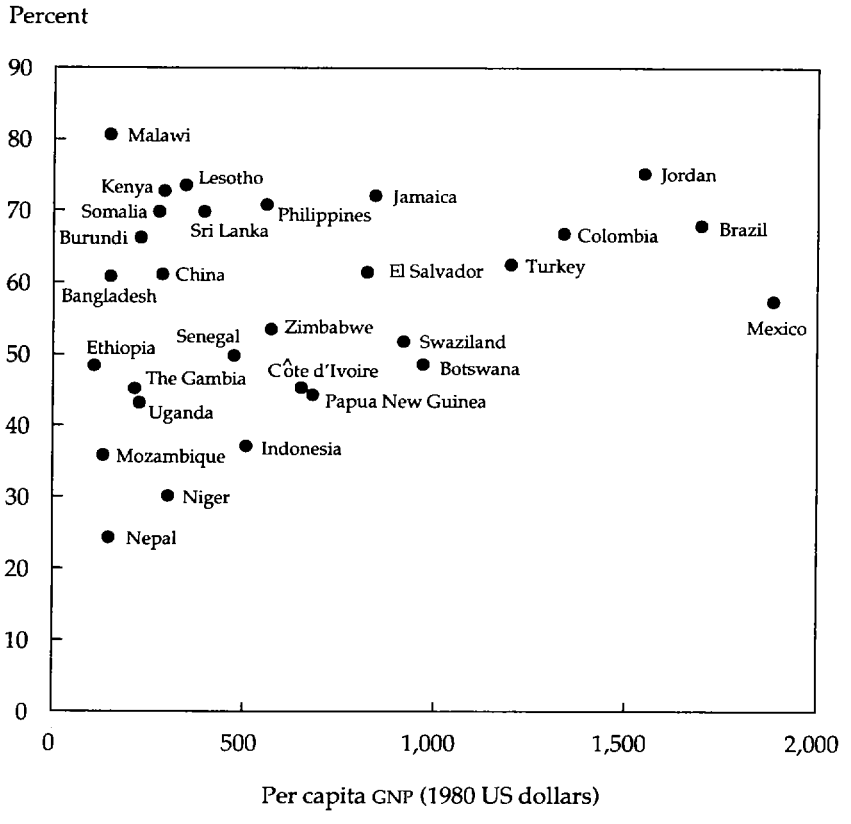


Source: World Bank data, 163 countries; most recent year, 1970–89.

of hospitals as high-cost institutions. A considerable share of the total health resources of a small country may be taken up by one large referral hospital because of that hospital's high cost and the country's limited resources. In Lesotho, for example, one hospital, the national referral facility, absorbed 42 percent of total Ministry of Health (MOH) expenditure in 1986–87 (United Medical Enterprises 1988). Reducing that share without significantly increasing the absolute level of resources in the health sector may imply drastically changing hospital functions or reducing government responsibility for its financing. Both of these actions are the subject of later chapters.

Some of the cross-country variation in hospital share may be attributable to differences in the definition of a hospital. In China, for example, if small township hospitals (which largely provide primary care) are not included with secondary and tertiary care facilities, the 1987 hospital share drops from 61 to 44 percent. The figures reported in table 2-1 reflect varied definitions of what constitutes a hospital in individual countries. This variation can mask the potential inequity and inefficiency resulting

Figure 2-2. Relation between the Share of Public Health Spending Devoted to Hospitals and per capita GNP, 1982–89



Source: World Bank sector reviews, project reports; see also table 2-1.

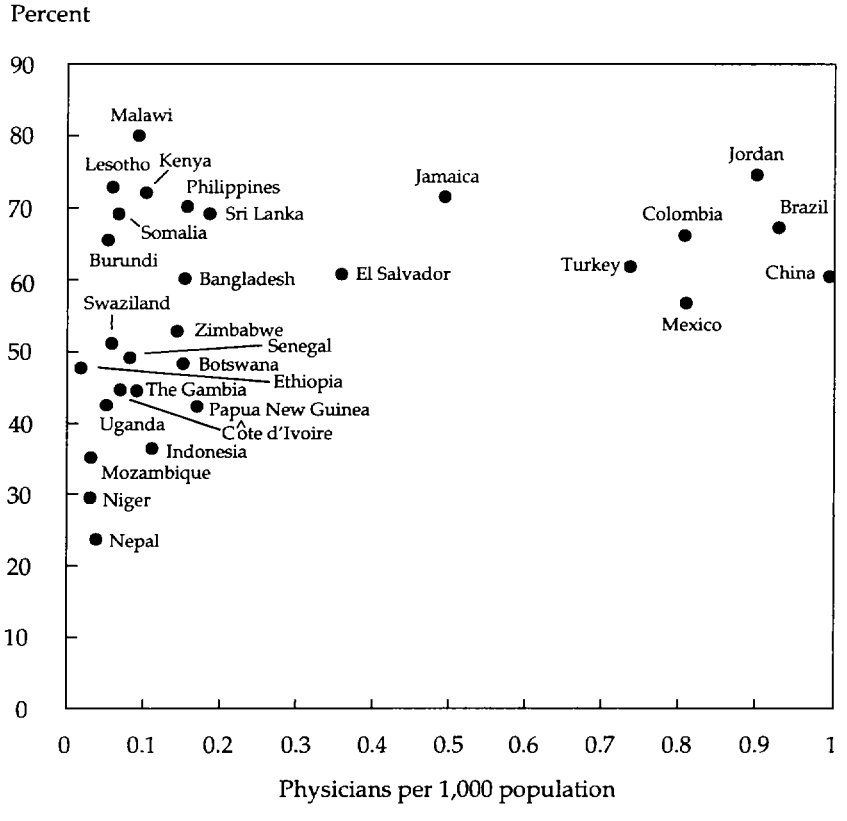
from a large proportion of resources being allocated to large urban tertiary hospitals as opposed to smaller district and rural hospitals.

Cross-country differences in the definitions of hospitals and in the share of recurrent hospital expenditures devoted to a few large tertiary hospitals compared with many smaller hospitals limit the usefulness of the overall public sector hospital share as an indicator of the priority given to various health programs. Comparing hospital with nonhospital expenditures is not equivalent to a comparison of primary care with secondary and tertiary care because hospitals provide a considerable amount of primary curative care. Therefore, as suggested above, the data for hospital share alone do not support definitive conclusions regarding

a country's level of support for primary health care or the effectiveness of primary, secondary, or tertiary care programs. Country-specific analysis is required for such conclusions.

Figure 2-3 indicates that there is a tendency for higher hospital shares to be associated with a greater number of doctors per capita. This relation is strongest for middle- and high-income countries. An examination of GNP per capita and doctors per capita as multiple determinants of the hospital share for twenty-nine low- to middle-income developing countries produces results that are not very strong but are consistent with this observation.[2] The statistical pattern in a cross-section of countries suggests that a 1.6 percent increase in the hospital share accompanies a 10 percent increase in the number of doctors per capita. Interestingly, the

Figure 2-3. Relation between the Share of Public Health Spending Devoted to Hospitals and Physicians per 1,000 Population



Source: See figure 2-2.

coefficient on per capita GNP becomes insignificant when per capita physician supply is included in the regression. A country's income level would still seem to be an important factor, however, because high-income countries tend to have greater availability of physicians per capita (see the appendix table A2-1).

In any case, the relation between the hospital share of health sector resources, the number of doctors per capita, and GNP per capita suggests that policies to limit the future supply of doctors (for example, redirecting subsidies for medical education to the training of other types of health professionals) might have an effect in the medium to long term on reducing the share of health resources absorbed by hospitals. Also, the correlation between physicians per capita and hospital share may indicate that the medical curriculum is oriented too much toward the practice of high-technology Western medicine and does not correspond to the health needs of the population. To a certain extent the transition in disease mix—away from the diseases of childhood and toward chronic disease and the diseases of adulthood—that accompanies growth in per capita income can require more hospital resources. A conclusion that is reached from a consideration of the cost-effectiveness of health sector alternatives later in the chapter is that primary health care services continue to have the greatest effect on health status per unit of expenditure even for countries that have advanced to middle-income levels.

Despite national and international calls for an emphasis on primary health care in the years since the 1978 World Health Conference held in Alma-Ata, the share of recurrent resources going to hospitals increased in many countries but remained about constant or decreased in other countries during recent years (table 2-2). The support that the data in the table lend to conclusions regarding a country's ability to reallocate resources is limited by the varying lengths of time between observations of the hospital share, but the data do suggest clear trends in some countries. The rate of annual increase of recurrent expenditures on hospitals compared with total health sector expenditures was 2.8 percent in Brazil (including nonhospital curative care), 6.6 percent in the Philippines, and 1.5 percent in Kenya, to pick some representative countries. In several other countries, the hospital share remained relatively constant, which may reflect the inertia of a historically based budgeting process (see Fiedler 1987 for a discussion of such a process in El Salvador). In some other countries, the hospital share decreased, usually as a result of deliberate government policies to control hospital expenditures and protect PHC programs. Mexico, The Gambia, and Zambia exemplify countries in which such policies have been translated into resource allocation decisions. In The Gambia, for example, the health plan adopted for the period 1981 to 1986 emphasized a village-based PHC strategy and described the shift in resources implied by this policy.

Table 2-2. Changes in the Share of Public Recurrent Health Expenditures Allocated to Hospitals, Selected Countries and Years
(percent)

Country	*Year*	*Hospital share*	*Average annual change*	*Source*
Botswana	1979	42		
	1984	49	3.0	
Brazil[a]	1975	70		
	1982	85	2.8	
China[b]	1980	38		
	1987	44	2.2	
Colombia	1974	65		
	1984	67	0.3	
El Salvador	1981	62		
	1985	62	-0.2	Fiedler 1987
The Gambia	1979–80	67		
	1985–86	45	-6.2	
Jamaica	1983–84	73		
	1986–87	72	-0.5	Kutzin 1989
Jordan	1982	80		
	1987	75	-1.2	
Kenya	1978–79	66		
	1985–86	73	1.5	
Lesotho	1970–71	64		
	1981–82	71	0.9	
Malawi	1976–77	73		
	1985–86	81	1.2	
Mexico	1982	64		
	1986	58	-2.6	
Papua New Guinea	1975	46		
	1987	45	-0.3	Newbrander 1987
Philippines	1981	55		
	1985	71	6.6	
Swaziland	1976–77	69		
	1983–84	52	-4.0	
Turkey	1981	68		
	1987	63	-1.4	
Zambia[c]	1975	33		
	1981	30	-0.5	

a. Percentages reflect the share absorbed by "curative" services (includes hospital and nonhospital).
b. Includes only upper-level hospitals and hospitals of traditional Chinese medicine.
c. Percentages reflect the share absorbed by the four largest hospitals only.
Source: World Bank sector reviews and appraisal reports, except as noted.

Although there were severe cuts in the real level of government health expenditure during the final three years of the plan period, the policy was adhered to, as evidenced by the data reported in table 2-2.

The recurrent resource needs of a country's health sector can interact dynamically with the decline of real public revenues that results, for example, from a fall in the prices of primary products coupled with softening economies in industrial countries and thus a drop in demand for these products. Nonhospital programs can therefore be placed in special jeopardy during times of economic difficulty, and a gap can be created between maintenance requirements and actual expenditures for hospitals. Declining public revenues have affected the quality and quantity of social sector outputs, although there has been no clear trend across countries as to whether these sectors have been affected disproportionately. Grosh (1990) found that, for eight countries from Latin America and the Caribbean between 1980 and 1988, declines in public spending on health exceeded declines in total government recurrent expenditures by an unweighted average of 25 percent. Public sector health expenditure, however, fell by less than total public spending in five of these eight countries, and total social sector spending (including education and social security) declined at only half the rate of total public spending.

The way the burden of the reduction in health revenue is shared across programs differs greatly among countries. The fear of many public health specialists is that, contrary to the example of The Gambia reported above, as the real value of revenue allocations for health services declines, health ministries, faced with a painful choice in allocation between hospital and nonhospital services, will favor hospitals, which have more visible and larger capital stock, over nonhospital programs, in which capital expenditures are less obvious and smaller. In the Philippines, for example, during a decline in real government expenditure between 1983 and 1985, the allocation to the Department of Health (DOH) was reduced to a greater extent than the overall decline in public spending. Within DOH expenditures, however, the share spent on hospitals increased, whereas expenditures on preventive programs declined rapidly (Intercare 1987). But some countries have reacted to decreasing health revenues by protecting or even increasing primary health care programs while reducing hospital recurrent expenditures. In Zambia, for example, a decline in copper prices put pressure on public resources available in the health sector in the late 1970s, and the share of tertiary hospital recurrent expenditures was reduced by 3 percent from 1975 to 1981 as the real value of health expenditures declined by 13 percent (World Bank 1984).

Investment

Hospital investment is by its nature lumpy and irregular, and the hospital capital stock produced by the investment has a life of two or more

decades. For this reason hospital investment expenditure for any given year can be a misleading indicator of the proportion of development funds absorbed by hospitals; the current size of the hospital sector results from much earlier health policies. Mills (1990a) contrasts selected data on the share of hospitals in investment in the 1960s with expenditures in the 1980s and notes the considerably higher share of capital expenditures spent on hospitals in the earlier period. For example, in Tanzania the hospital share of total health sector investment was reduced from 80 percent in 1961–62 to less than 32 percent in 1978–79. The high early figures result from large-scale hospital construction programs started in the 1960s, whereas the recent lower figures reflect the more limited donor funds now available for hospital construction and the greater priority given to primary health care.

One cause of imbalance between recurrent resource availability and capital expenditures is that, in many countries, health ministries are largely responsible for recurrent budgets, whereas donors fund a large proportion of capital expenditures. Somalia provides an example of a very poor country in which both the recurrent cost and the investment cost of primary health care programs have been heavily supported by donors, but hospital recurrent costs are left to be supported almost strictly from government revenues (World Bank 1985). There is evidence that in recent years some health ministries have favored hospitals in their marginal recurrent resource allocation decisions and have focused on primary care in their capital investment decisions, which are largely donor-determined. Underfunding of the recurrent resource requirements of these projects may in part be the result of the unwillingness of health ministries to adjust their marginal recurrent resource allocation decisions from an earlier hospital bias. Kenya provides an example of public sector health authorities' unwillingness or inability to shift recurrent resources in favor of primary care investments. A comparison of Ministry of Health recurrent budgets with actual expenditures from 1982 to 1989 reveals that for curative services (mostly hospitals), annual budgets were overspent by an average of 7 percent and actual allocations to preventive and promotive programs and rural health services were 17 percent and 19 percent less, respectively, than the amounts budgeted. This evidence of actual resource allocation runs counter to the ministry's stated policy, as reflected in MOH budgets, of giving priority to rural, preventive, and promotive health services (World Bank 1991). A similar pattern existed in Jamaica from 1983–84 to 1986–87; expenditures on hospital services were consistently above the amount budgeted (by an average of 16 percent for these four years), whereas primary health care expenditure averaged slightly less than the amount budgeted (Kutzin 1989).

Hospital investment and recurrent expenditures are linked closely but with substantial lags. Present recurrent expenditures on hospitals result

from the cumulative hospital capital stock produced in the 1960s and 1970s and in many cases, especially in Africa, during an earlier colonial era. Consistent procedures are needed to assist planners in projecting the recurrent costs of project proposals and then incorporating these projections into alternative budget scenarios. Ultimately, a careful project analysis by component and expenditure category needs to be made. Recurrent–capital cost (RCC) ratios provide a convenient alternative, however, prior to a more detailed analysis. Projections of recurrent expenditures based on RCC ratios are not reliable for a detailed analysis of the recurrent cost implications of specific investments because the relation between variable costs, fixed costs, and outputs that underlies the ratios may not be stable across countries, projects, and levels of output. Nevertheless, RCC ratios can be a valuable method of making quick first estimates of the recurrent costs associated with a contemplated investment strategy. It is instructive to compare the recurrent cost ratios reconstructed from a number of more detailed sources. Using accounting-based analyses of existing hospitals or analyses of new projects to estimate recurrent costs, the studies summarized in table 2-3 show RCC ratios varying from 0.10 to 0.40 and averaging roughly 0.20. Some of the cross-country variation results from methodological differences described in the "Comments" column of the table (for example, inclusion of depreciation in recurrent costs), and some probably results from differences in budgeting conventions. Despite what may be characterized as these artificial sources of cost variation, most of the estimated RCC ratios in table 2-3 are clustered around the mean value of 0.20. These ratios are substantially lower than for projects that are less capital intensive, such as primary health care outreach programs, which can entail ratios in excess of 0.5 (Heller 1979; Over 1981).

Allocation of Resources by Level of Hospital

The scope and complexity of hospital services vary greatly within countries. To sharpen the discussion it is necessary to categorize hospitals by the level of services offered.

Definition of hospital levels. Differences in case mix, technical capacity, and skills differentiate hospital levels. These differences may also imply different sizes of facilities, roughly measured by the number of operational beds, upper-level facilities often having a larger size. Bed size is not the definitive difference between levels of facilities, however, and there are many examples of facilities in which bed sizes run counter to the expectation of size based on level. It is difficult to distinguish between levels of facilities when one set of criteria is used for all countries. For convenience, three levels of general hospital facilities are distinguished here, not on the basis of bed size but on the basis of service characteristics.

Table 2-3. Recurrent Cost Ratios for Selected Hospital Projects

Country	*Level of hospital*	*Year*	*RCC ratio*	*Source*	*Comments*
China	Central Provincial District	1986	0.22 0.22 0.40	Barnum 1989	Comparison of reported operating costs and replacement cost.
Indonesia	Central Provincial District	1986	0.16 0.12 0.12	Barnum 1987	Based on accounting analyses for existing hospitals compared with new project costs.
Jamaica	District	1989	0.16–0.18	Kutzin 1989	Based on accounting analysis for new project.
Kenya	Unspecified	c. 1970	0.10–0.30	Heller 1979	
Malawi	District	1988	0.14	Mills 1991	Based on accounting analyses for existing hospital compared with replacement cost.
Malaysia	Provincial District	1971	0.18 0.11–0.30	Heller 1975	
Papua New Guinea	Central Base Provincial	1988	0.27 0.27–0.40 0.22–0.33	JSI 1990	Comparison of recurrent costs (including depreciation) and estimated capital replacement cost.
Rwanda	Central Provincial District	1987	0.18 0.13 0.14		Based on studies of similar facilities plus data gathered locally.
Uganda	Unspecified	1970	0.25–0.30	Dunlop 1973	
Zambia	Central Provincial	1981	0.30–0.40 0.16		MOH accounts compared with estimated capital replacement cost.

Source: World Bank sector reviews and appraisal reports, except as noted in table.

The classification fits a pyramidal conception of a health service referral system, ranging from level I, the most technically complex, to level III, the most basic.[3] The levels can be described as follows:

Level I: These hospitals have the most specialized staff and technical equipment. The clinical services are highly differentiated by function, including, for example, cardiology and specialized imaging units. More than ten clinical specialties are not unusual in large referral centers. The staff-to-bed ratios are the highest in relation to other hospitals. Hospitals at this level are often, but not exclusively, associated with a medical school. As befitting their location at the top of the referral pyramid, level I hospitals are also called "central"- or "tertiary"-level hospitals. Bed size, as suggested above, varies; depending on the size of the population to be covered, it ranges from 300 to more than 1,500 beds.

Level II: Although lacking the most technical services available in level I hospitals, level II hospitals are also highly differentiated by function, with five to ten clinical specialties. Staff-to-bed ratios are intermediate between the highest and lowest level of hospitals. Level II hospitals are occasionally associated with medical schools. Bed size ranges from 200 to 800 beds. Level II hospitals are often referred to as "provincial" hospitals.

Level III: Level III hospitals have many fewer specialists than the higher-level hospitals. The availability of skill levels within this category is varied. Some entry-level hospitals may have specialists in internal medicine, obstetrics-gynecology, pediatrics, surgery, or radiology, whereas others may have only general practitioners. Limited laboratory services are available for general but not specialized pathological analysis. Staff-to-bed ratios are low in relation to the upper-level hospitals. Level III hospitals are often referred to as "district" or "first-level referral" hospitals.

These definitions are by necessity very approximate and meant only to be suggestive. The actual differences between referral levels depend greatly on the resources available, the extent of health sector development, and the standards of service within a given health system. Additional levels could conceivably be distinguished. Level III is particularly large and, for example, includes both class "C" and class "D" hospitals in Indonesia, "county" and "township" hospitals in China, and district hospitals in Malawi.

Hospital level and resource use. Higher-level (level I and II) hospitals absorb a large share of total public expenditure on hospitals, despite the small number of these facilities in comparison with the number of level

III hospitals in most countries. The greater case mix complexity in higher-level hospitals and their more intensive input use commonly translate into much higher operating costs per unit (see chapter 3) than at lower levels. Hospitals that have a teaching function also tend to have higher operating costs. This difference is expected on the basis of function, and thus it is expected that level I and II hospitals will absorb a greater proportion of total hospital resources than simply their numbers would indicate. The extent to which resources are concentrated in these facilities, however, often goes beyond that which might be justified in order to fulfill their tertiary functions. In Zambia, for instance, the three large central hospitals use 30 percent of Ministry of Health resources and an estimated 45 percent of total MOH hospital resources, leaving the remaining 55 percent to cover thirty-nine lower-level hospitals (World Bank 1984). Similarly in Indonesia (Barnum 1987), the nineteen largest public hospitals (out of more than three hundred) used 53 percent of total recurrent hospital expenditure in 1985. In Kenya (World Bank 1991), Kenyatta National Hospital used almost 20 percent of recurrent MOH expenditure for curative services in 1986–87. Belize City Hospital in Belize (Raymond and others 1987) used 69 percent of recurrent MOH hospital expenditure in 1985. Finally, in Zimbabwe (Hecht 1992), 45 percent of MOH recurrent hospital expenditure in 1987 was for four central hospitals.

This concentration of resources may have undesirable implications for overall sectoral resource use and equity. In each of these cases a relatively modest proportion of total hospital expenditures remained to finance the services of district hospitals, which are intended to serve as first-level referral facilities for most of the population. Excessive centralization of resources in a small number of tertiary hospitals is not consistent with the performance of crucial first-line functions by district hospitals, because these smaller hospitals are relatively starved of inputs. Such centralization may result in poor-quality services at these first-level referral centers, possibly causing patients to bypass them in favor of the tertiary hospital. The end result is inefficiency in the overall network of health services. The importance of maintaining a reliable supply of services with an acceptable level of quality to enable district hospitals to perform their first-level referral role is examined in greater detail in chapter 6.

Distribution of Resources within Hospitals

The distribution of hospital expenditures by line item category reflects the input mix used in the production of services. Depending on management incentives, hospital input mix may be determined by many influences in addition to input prices, including centrally determined

personnel staffing quotas and nationally acceptable medical practices. In any case, relative prices will strongly influence the share of expenditure on different inputs. In more industrialized countries, labor inputs (wages and salaries) commonly absorb a high proportion of operating expenditures, because wages and salaries tend to be high in relation to nonlabor inputs (Mills 1987). In contrast, in developing countries salaries and wages tend to be lower in relation to nonlabor inputs, and thus it might be expected that labor cost would be a smaller proportion of total expenditures. This expectation could be disappointed to the extent that economic incentives lead to substitution of equipment and other nonlabor inputs for labor in high-wage countries and the reverse in low-wage countries, but accepted standards of care and medical procedures provide limits on this substitution.

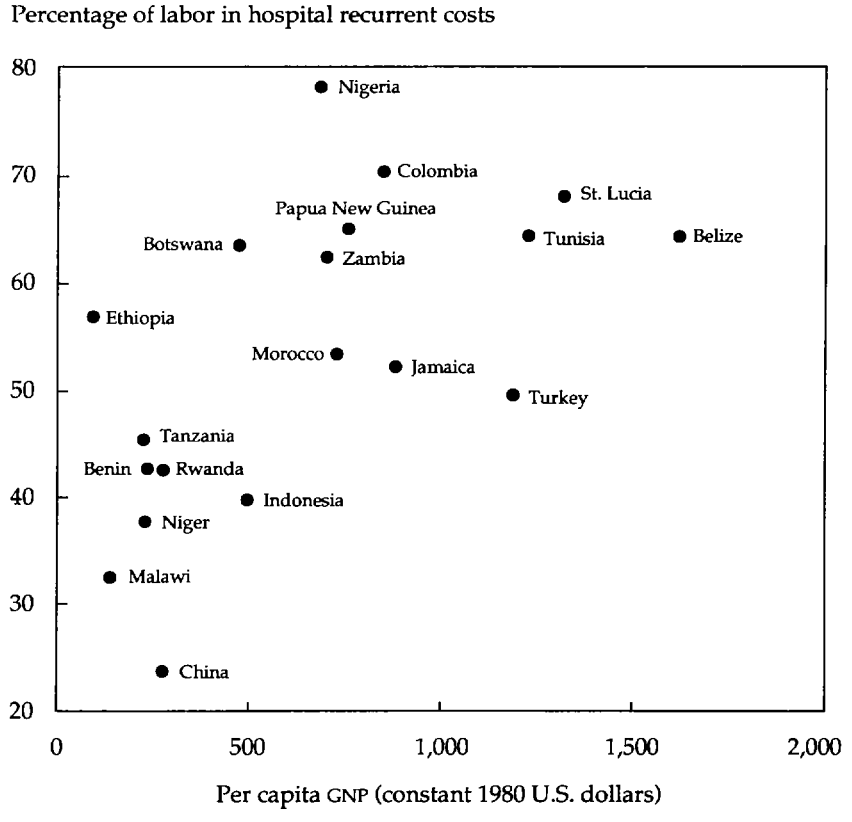
Data on labor shares for a selection of hospitals in a cross-section of countries support the expectation that the share of labor increases with income, at least up to the middle-income levels indicated in figure 2-4, in which the percentage of labor in hospital recurrent costs is compared with GNP per capita.[4] Limits on the substitution of nonlabor inputs are also created by rigidities in employment policies and the setting of wage rates. The share of total hospital expenditure devoted to personnel has increased in many countries as a result of a decreased availability of public sector revenues and an inability or unwillingness to reduce the size of hospital staff, possibly because of civil service constraints on the hiring and firing of staff.

Even though in low-income countries the share of total expenditures absorbed by labor tends to be less, labor remains an important input. The data in table 2-4 show that in African countries labor costs are 33 to 79 percent of total expenditures. In Indonesian hospitals, personnel costs are about 40 percent of total expenditures, and in Jamaica and Belize, personnel costs range from about 50 percent to 74 percent of total hospital expenditures.

In China labor costs are 23 to 26 percent of total expenditures. In part, this unusually low personnel share may be caused by the exclusion of the value of nonmonetary or budgeted benefits from hospital accounts. Still, the low personnel share also results from a deliberate policy to emphasize nonlabor inputs, especially pharmaceuticals, in inpatient and outpatient hospital services. This latter factor is related to China's method of financing hospitals through user charges; drugs are priced at profitable levels, thus encouraging their use by providers. Largely because of this situation, drug fees are the leading source of hospital revenues.

Drugs and medical supplies vary greatly as a percentage of total hospital expenditures. These are often largely acquired in international markets with foreign exchange and at internationally determined prices.

Figure 2-4. Relation between the Share of Labor in Hospital Recurrent Costs and per Capita GNP, 1978–90



Source: See table 2-4.

Thus the expenditure on drugs and medical supplies does not reflect local prices to the same extent as labor costs. If the relative quantities of drugs and labor used per bed-day in countries with low wages were similar to those used in higher-wage countries, the share of expenditures on drugs would be higher. Central ministry pressures (in addition to relative prices) may dictate, however, that a lower quantity of medical supplies be used per unit of labor in low-wage countries. This may be a rational response to prevailing relative prices, providing allocational efficiency (as at point A in figure 2-5, where total output is 1,000 bed-days), or it may be a suboptimal response, forcing an economically

Table 2-4. Hospital Recurrent Costs by Line Item Category
(percent)

Country	Year	Level of hospital	Total recurrent cost: Personnel	Drugs	Other	Total	Number of hospitals in study	Source
Belize[a]	1985	II	56	—	44	100	1	Raymond and others 1987
		III	74	13	13	100	6	
Benin[b]	1986	II	43	13	44	100	1	Pangu and others 1987
Botswana	1978	All	64	17	19	100	All MOH	BMOH 1979
Colombia[b]	1978	I–II	71	15	14	100	8	PRIDES 1980
China	1986	I	23	38	40	100	8	Barnum 1989
		II	26	43	32	100	11	
		III	24	50	26	100	7	
Ethiopia	1983–84	Urban	56	17	27	100	9	Donaldson and Dunlop 1987
		Rural	59	17	24	100	11	
Indonesia	1985	I	33	—	67	100	2	Barnum 1987
		II	42	—	58	100	15	
		III	39	—	61	100	296	
Jamaica	1985–86	I	55	10	36	100	2	Kutzin 1989
		II	49	12	39	100	2	
		III	55	11	34	100	1	
Malawi[c]	1987–88	III	33	30	37	100	6	Mills 1991
Morocco	1987	All	54	24	22	100	All public	Bennis and others 1990
Niger	1986	I	38	34	28	100	1	Wong 1989
Nigeria	1986	III	79	16	5	100	2	
Papua New Guinea	1988	I	73	9	18	100	1	JSI 1990
		II	63	9	28	100	4	
		III	64	9	27	100	8	

(Table continues on the following page.)

Table 2-4 (continued)

Country	Year	Level of hospital	Total recurrent cost				Number of hospitals in study	Source
			Personnel	Drugs	Other	Total		
Rwanda	1987	I	43	—	57	100	1	Shepard 1988
St. Lucia	1986	II	69	8	23	100	1	Russell, Gwynne, and Trisolini 1988
Tanzania	1979	All	46	35	19	100	All MOH	Dunlop 1984
Tunisia[d]	1990	I	65	12	23	100	22	
Turkey[b]	1987	I	50	30	20	100	3	
Uganda	1968	All	63	14	23	100	All MOH	Dunlop 1973
Zambia	1981	I	60	—	40	100	3	
		II, III	63	—	37	100	All MOH	

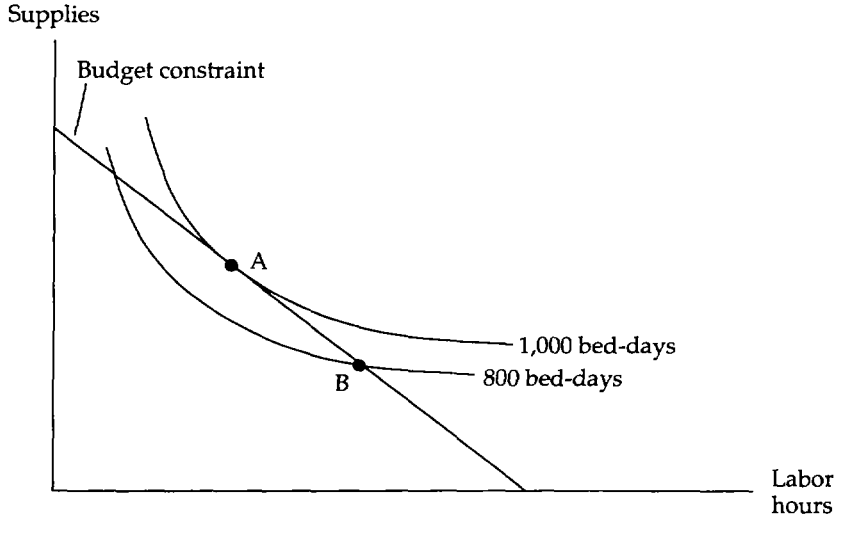
a. "Drugs" includes medical and nonmedical supplies.
b. "Drugs" includes all medical supplies.
c. Total amount includes some nonhospital district health expenditures.
d. Figures reflect amount budgeted for a combination of teaching hospitals and specialized institutes.
— Not available.
Source: World Bank sector reviews and appraisal reports, except as noted in table.

inefficient solution (as at point B in figure 2-5, where output is 800 bed-days even though the budget remains constant).

In poorer countries, drug shortages are common and absolute expenditures on drugs are limited, so that even though wages are low, drugs as a percentage of total costs may not be high. Many analysts have argued that the shortage of drugs forces technical inefficiency on the hospital sector. This is particularly true if rigid personnel policies make substitution of supplies for personnel difficult. It is the combination of drug and supply shortages (and other manifestations of recurrent cost constraints, such as insufficient maintenance) and rigid personnel policies that creates the technical inefficiency. Improved hospital information systems coupled with management policies and training could lead to greater flexibility by central authorities, hospital administrators, and physician-managers with respect to input substitution. Such reforms would be positive steps toward reducing this source of suboptimal hospital performance. The potential gains are limited, however, by operational constraints on managerial choice of inputs that may be unaffected by these reforms (lack of drugs or equipment, for example).

One possibility for improving the internal efficiency of hospitals may be to change the mix of labor inputs in the production of hospital

Figure 2-5. Underfunding of Supplies



services. Many hospitals in developing countries are organized and staffed based on industrialized country models. The different resource endowments and epidemiological profile of low-income countries suggest that such a model may not be appropriate here. Very little information is available on hospital staff mix in developing countries, and microeconomic-level studies are needed to examine the scope for improving the cost-effectiveness of hospital services through the reallocation of staff and the creation of new staff categories. This is a sensitive issue because medical (and in some cases, nursing and paramedical) professionals may perceive such reallocation as a threat to the positions they have worked hard to attain. Also, the internal organization of large hospitals (Harris 1977) leads to a bifurcation of labor between hospital support services (for example, laundry, pharmacy, blood bank) and direct medical care. Input mix and labor allocation on the medical side may be viewed as determined by scientific standards. These standards, however, are often determined in industrial countries. The medical input mix should not be viewed as fixed (Harris 1977). Training and operational research can improve staff flexibility and increase medical decisionmakers' awareness of substitution possibilities. Consideration of alternative staffing patterns is particularly critical for those countries that are suffering severe shortages of skilled staff. Some substitution examples follow:

- *Substitution of nonphysician- for physician-managers.* Physicians in hospitals commonly perform management functions that could be more effectively and cheaply filled by trained managers who are not physicians. The use of professional hospital managers is common in industrial countries but rare in middle- and low-income countries. The principal reason for this is the strong tradition in the medical profession of physician control of hospitals. In many countries the lack of appropriate training and internship programs is a severe constraint on the development of a cadre of career hospital managers.
- *Substitution of nursing for physician time.* Wider use of nursing staff could reduce the use of physician time in the hospital. Monitoring of patient condition, basic diagnosis, obstetrics, and simple curative intervention for trauma are possible areas in which nursing training and activities could be extended.
- *Substitution of clerical staff for registered nurse time.* In some countries, registered nurses (RNs) are widely used for clerical tasks and in some cases (primarily in Africa) for nontechnical nursing activities. Many of the tasks performed by RNs there could be adequately carried out by less-well-trained persons. Clerks could perform some clerical functions, and assistant nurses could perform some patient care and other

ward-related activities. The use of alternative staff could free RNs to focus their efforts on activities that make the best use of their skills.

Distributional Equity of Hospital Use

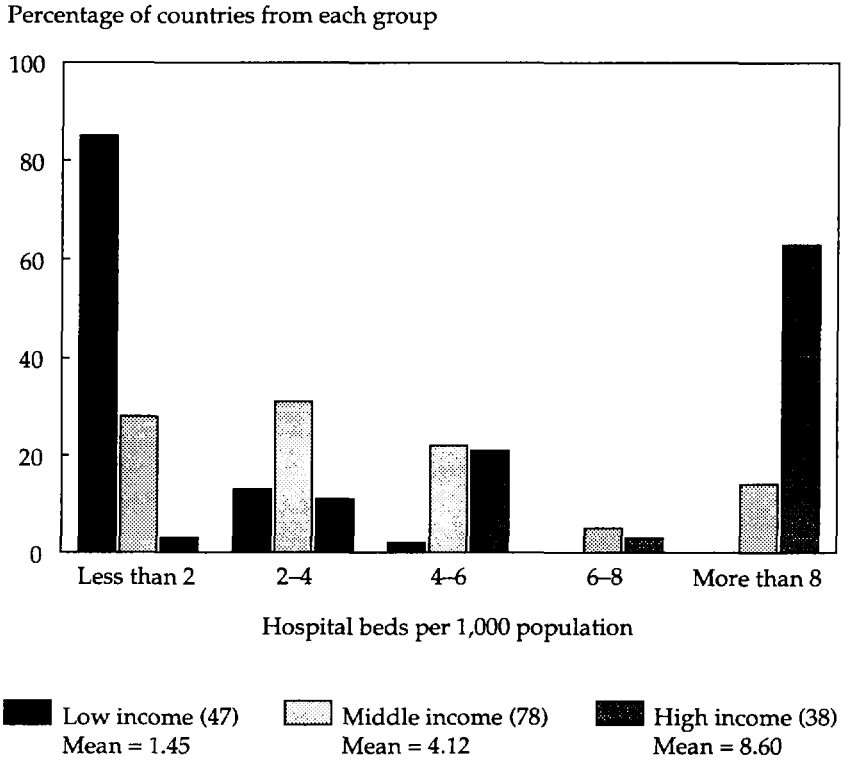
With a cross-country mean of about 60 percent, hospitals absorb a substantial percentage of public recurrent health expenditures. The share of health resources devoted to hospital services is believed to be inversely related to the equity of overall health service provision; the greater the share devoted to hospitals, the less is left for primary care programs and facilities that have the potential to provide relatively low-cost basic curative and preventive services to a broadly dispersed population. Thus, within the health sector, the share of resources absorbed by hospitals has implications for the equity of the service delivery system.

The characteristics of hospital services themselves may have a profound effect on equity in many poorer countries in addition to the effect the share of government health resources devoted to hospitals may have on the availability of primary care services. Hospitals use a large part of the total public sector health budget but can provide benefits to relatively few. Even with existing large expenditures only a limited number of beds can be provided, and inpatient service remains necessarily restricted to a small number of people. Hospital bed ratios per 1,000 population in low-income countries are commonly less than two, in contrast to means of more than four beds per 1,000 in middle-income and more than eight beds per 1,000 in high-income countries (see figure 2-6). In this section we describe the distribution of hospital expenditure and use across income groups, geographic location, age, and disease categories. In addition, we assess implications of projected demographic and epidemiological changes for the distribution of hospital services among population subgroups.

Hospital Use across Income Groups

Little recent research has been conducted on the income characteristics of hospital users. However, the analyses of individual hospitals in several Asian countries suggest that hospitals are used unequally by different socioeconomic groups within designated catchment areas. A survey of inpatient income from a sample of seventeen hospitals in Malaysia showed that only 21 percent of hospital inpatients were from households below a specified poverty criterion compared with 39 percent of the general population (Heller 1975). Upper-income groups used a substantially greater share of total inpatient services than would be expected from their proportion of the population. Using 1978 data from the Bicol

Figure 2-6. Distribution of Hospital Beds per 1,000 Population in 163 Countries, Most Recent Year, 1970–89



Source: World Bank data.

region of the Philippines to examine the distribution of public health subsidies across income groups, Ching (1986) found that, for public hospitals, the per capita subsidies to the poorest three quintiles of the population were below the regional mean, whereas those to the wealthiest two quintiles were above the regional mean. In fact, per capita subsidies were greatest to families in the richest quintile. A household survey in Indonesia revealed that households in the lower 40 percent of per capita income distribution accounted for only 16 percent of hospital service use, whereas households in the upper 30 percent accounted for 63 percent (Meesok 1984). Hospital-based surveys conducted in 1985 in the Nusa Tenggara Barat province of Indonesia had similar findings.

Patients from the richest 9 percent of the province's income distribution accounted for 32 percent of inpatients and 56 percent of outpatients. Conversely, the poorest 55 percent of the population accounted for only 32 percent of inpatients and 18 percent of hospital outpatients (Gish, Malik, and Sudharto 1988). Thus, persons in the highest income groups benefited disproportionately from Indonesia's subsidies to public hospitals, with a greater degree of regressivity involved in the use of hospital outpatient care.

A study of the characteristics of patients at one African tertiary hospital had results similar to these Asian studies. Surveys of patients at Niamey National Hospital in Niger showed that inpatients had a median income that was comparable to or slightly higher than other urban residents (who had, in turn, higher incomes than rural residents), and outpatients had a higher median income than inpatients (Weaver, Handou, and Mohamed 1990b). The pattern suggested by this and the above studies is that government subsidies to public hospitals are not well targeted to the poor and that this regressivity is magnified for hospital outpatient services, perhaps because of their more discretionary nature.

Not all the available information (particularly studies from Latin American countries), however, confirms inequitable use across income groups. Selowsky (1979) used household survey data in conjunction with public sector and social security expenditures on health services to analyze the distribution of public subsidies for health across income groups, facilities, and regions in Colombia in 1974. He found that, in public sector hospitals, the lowest-income quintile received the greatest subsidy for inpatient care and that this share declined monotonically for higher-income groups. McGreevey (1990) cites information for five South American countries (Argentina [1980], Costa Rica [1982], Chile [1982], the Dominican Republic [1980], Uruguay [1982]) that demonstrates that the benefits from social security expenditures on health, much of which is used for public hospital services, are inversely related to income. In those few countries in which social security covers most of the population, as in Brazil, the overall distributional effect of public hospital services is to improve equity. One possible explanation for the better equity effects of public hospital subsidies in Latin American countries is the existence of a strong modern private sector serving the wealthiest population groups, leaving lower-income groups to use the public (ministry of health or social security) hospitals.

Similar studies in other countries are needed, especially ones distinguishing between the use of public and the use of private hospital services by socioeconomic group. When there are perceived quality differences between public and private services, either material differences related to health outcome or differences in amenities, high-income

groups will favor private services. In Sri Lanka, for example, the value of per capita use of public services by low-income households far exceeds the expenditure by such households on private care, whereas the value of government services received by the highest-income group is about one-half that group's expenditure on private medical care (Alailima and Mohideen 1984).

Geographical Distribution

Hospitals are located primarily in urban areas, and even when they are intended to provide a referral service for a broad geographical population base they serve in actuality a disproportionately urban clientele. Because urban populations generally have higher incomes than those of rural areas, the urban location of most hospitals affects income equity as well as geographic equity. The urban bias provided by hospital services extends even to China, which has decreased mortality in rural areas during the last four decades and is widely perceived to have focused on equity. Other factors, such as improved nutrition and education, have probably been at the root of China's improved health status in this period. In a study of the distribution of health care resources, Prescott and Jamison (1985) noted that the fall in mortality has not been achieved through a more equitable distribution of health resources. There is a wide variation in the availability and distribution of hospital beds, health workers, and health expenditure among Chinese provinces. They found that the distribution of health resources is strongly related to the degree of urbanization in the province and to urban income. In their study, a 1.0 percent increase in urban income is associated with a 1.5 percent increase in the ratio of medical doctors to population, a 2.5 percent increase in the general hospital bed ratio, and a 3.0 percent increase in recurrent expenditure. These increases are also likely correlated with the extent of health insurance coverage, which is associated with greater health care use and expenditure, as well as with greater income.

Similar inequalities in the availability of hospital beds and in the distribution of hospital expenditures are documented in health sector assessments of other countries. Evidence from a few representative studies can be cited. The number of hospital beds per thousand population in Indonesia varies from 0.2 to 1.2 across provinces, and the variation again is related to levels of urbanization and income (Barnum 1987); the income elasticity of hospital beds is 1.0 (Prescott 1991) and, given that hospital beds located in more heavily urbanized and high-income areas tend to be staffed and equipped more intensively, the income elasticity of hospital expenditure is even greater.[5] In Malaysia (Heller 1975) and Papua New Guinea (Baker 1977), hospital expenditure per capita varied more than twentyfold across provinces. In Brazil, expenditure per capita

by the social security health system in the lower-income northeastern region in 1986 was less than half of those of the higher-income states in the southeast (World Bank 1988), and the availability per capita of hospital beds in the north and northeast was about half of that in the south and southeast in 1984 (Briscoe 1990). In Colombia the average government health subsidy per urban household in 1974 was more than twice as large as that received by rural households (Selowsky 1979). In many African countries probably at least equal variation exists because of the predominantly urban location of hospitals and the largely rural, dispersed populations.

The above examples are not intended to constitute an argument for establishing hospitals in rural areas to meet equity concerns but to highlight the importance of improving the reliability of lower-level facilities that serve more widely dispersed populations. The examples also draw attention to the importance of improving transport and, more generally, the referral network to widen access to hospital services. The "bias" of hospital services toward urban dwellers is to be expected from the nature of hospitals as institutions with high fixed costs. It would not make sense to establish large hospitals in sparsely populated areas or rural areas with a widely dispersed population; a network of low-cost primary care that included smaller district hospitals would seem to be more appropriate for these areas. Other risks are associated with using hospital investments in poor geographic areas as a means of improving equity, such as the creation of a two-tiered medical system, in which poor quality and stigma are associated with the facilities serving poor communities. It is logical for hospitals to be located in areas with a concentration of population high enough to keep the relatively large amount of resources (staff, equipment, supplies) occupied. Ensuring that those not living near these areas have access to hospital care when needed is the function of the referral system.

In concept, referral within the health system is supposed to provide a wider distribution and appropriate use of health services, entry being gained at the bottom of the referral pyramid through primary care facilities and other providers. In practice, self-referral—often resulting from the perception or fact of poor quality primary care services because of a lack of drugs or personnel—and poor administrative, transport, and communications linkages between levels of care, frequently defeat the intended purpose of the referral system. Many referral studies have revealed that hospital entry is a matter of proximity, rather than the appropriateness of the classification, and that most inpatients in upper-level hospitals in the most urbanized areas are admitted directly. This situation is to be expected in the absence of barriers to easy admission, when, for many urban dwellers, the nearby hospital is the closest source of primary care. A sample of hospitals in Indonesia revealed that more

than 50 percent of inpatients come from less than 5 kilometers away from the hospital and more than 80 percent of outpatients come from less than 1 kilometer away (Barnum 1987). Cumper, Walker, and MacCormack (1985) concluded that much of the variation in hospital admission rates across census constituencies in Jamaica (from 13 to 92 per 1,000) could be explained by hospital location. The referral system is discussed in greater detail in chapter 6.

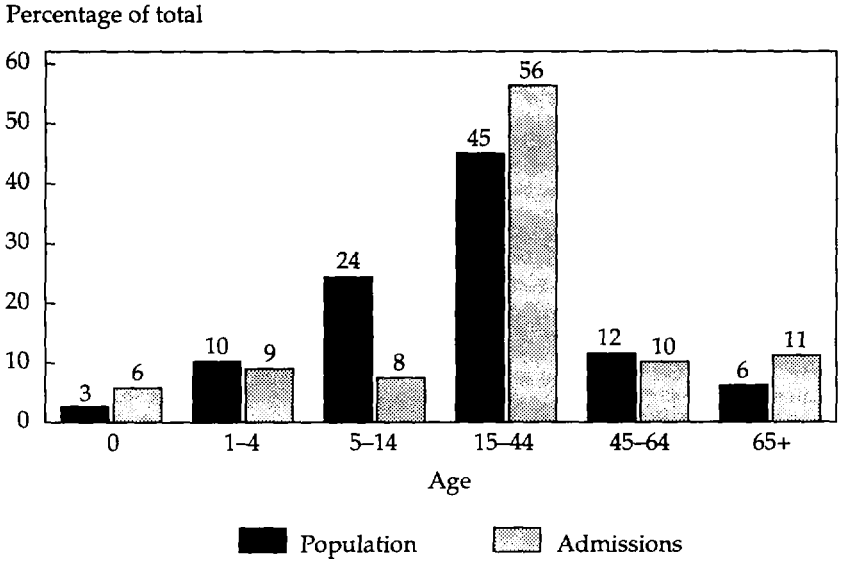
Distribution of Hospital Use by Age

Both the overall level of demand for hospital services and the specific services required are affected by the age and disease profiles of the population. As the economic development of a country proceeds, both profiles change in what is commonly characterized as demographic and epidemiological transitions. The demographic transition refers to the change in the population age structure that accompanies the fall in fertility and mortality as economic and social development occur. The epidemiological transition refers to the changes in age-specific death and illness rates that result in reduced infant and child mortality and longer expected life span. Interaction of these two transitions results in an aging population that is characterized by a lower prevalence of infectious disease and a higher prevalence of chronic disease (Jamison and Mosley 1990). The implications of this transition for hospital resource use can be derived from the experience of specific countries.

Hospital inpatient services are used primarily by adults, both because they comprise the largest proportion of the population and because children, after they survive their first year of life, are much less likely to need hospitalization (for epidemiological reasons) than they will in later life. For example, according to samples from hospitals in the countries concerned, adults comprise 88 percent of hospital admissions in county hospitals in China (Over and others, 1992); 87 percent in Papua New Guinea (JSI, 1990); 84 percent in Belize (Raymond and others 1987); and approximately 70 percent in Malawi (Mills 1991), Niger (Wong 1989), and Uganda (Over and others, 1992). The lower percentages of adult use of hospitals in the African countries reflect the smaller percentage of population that is more than fifteen years of age. Significantly, however, it also reflects the reduction in child mortality and the increase in the relative importance of chronic diseases in China and Papua New Guinea during the last twenty years.

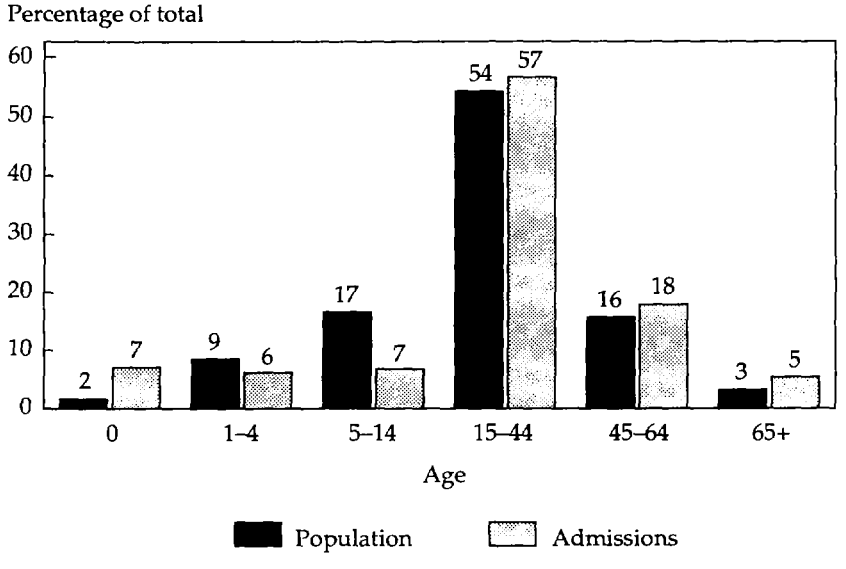
This pattern of use is apparent in more detail in figures 2-7a and 2-7b, which present comparisons of the age distribution of the population with the age distribution of hospital admissions in Jamaica and the Republic of Korea.[6] The figures indicate that, in both countries, the distribution of hospital admissions by age corresponds closely with the

Figure 2-7a. Distribution of Population and Admissions by Age Group, Jamaica, 1985



Source: GOJMOH 1987.

Figure 2-7b. Distribution of Population and Admissions by Age Group, Korea, 1986

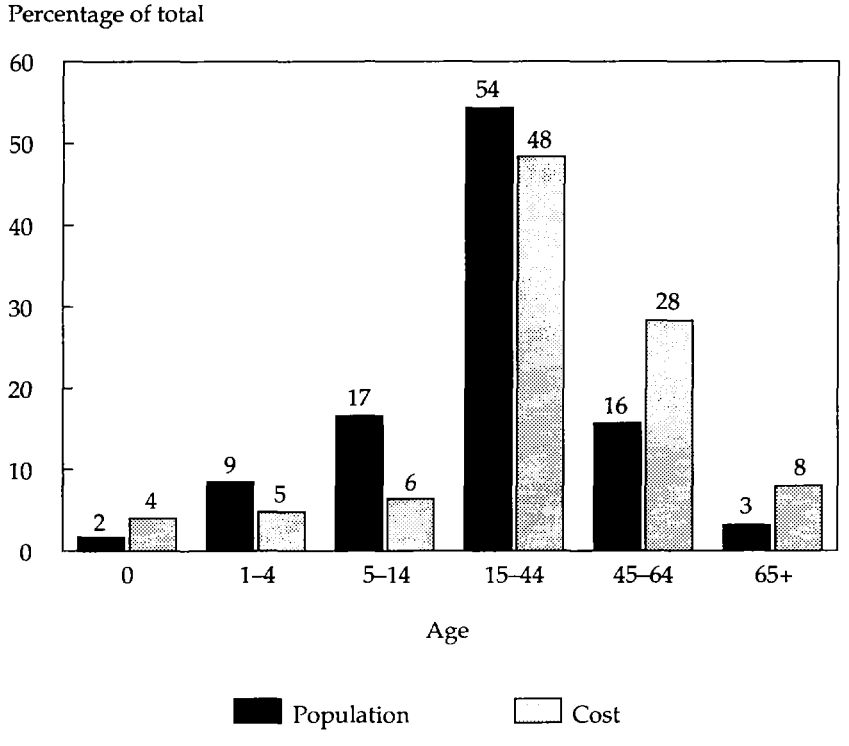


Source: FKMIS 1987.

age distribution. Exceptions are the group age five through fourteen, which makes least use of hospitals compared with that group's share of the population, and infants (less than one year), who make most use of hospitals in relation to their population share. The elderly also tend to use hospitals disproportionately, as would be expected.

The wealth of data available on insured patients from Korea provides a better look at inpatient resource use by age group. In figure 2-8 the population distribution of the insured is compared with the age distribution of inpatient expenditure. In Korea, most inpatient care is provided by private hospitals that are financed through a combination of insurance reimbursement, copayments, and user fees. Therefore, expenditure data provide a good proxy for measuring resource use. Conclu-

Figure 2-8. Population and Inpatient Cost for Insured Patients by Age Group, Korea, 1986



Source: FKMIS 1987.

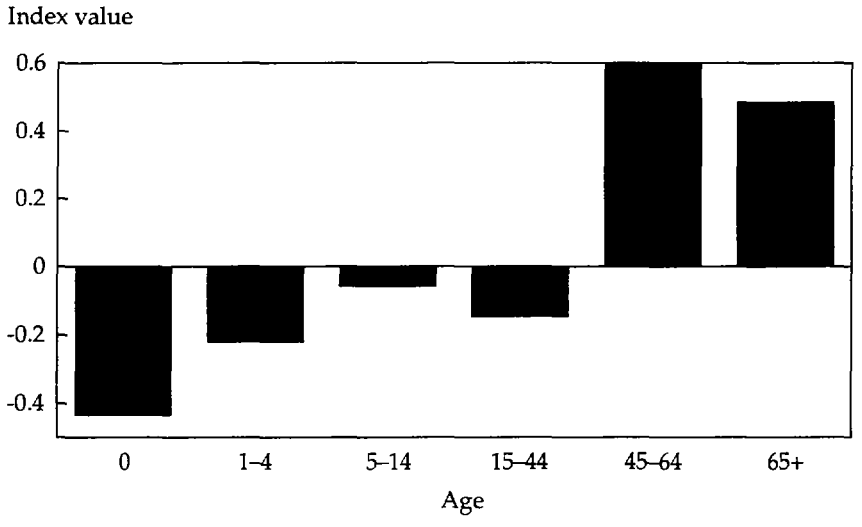
sions drawn from this figure are similar to those from figures 2-7a and 2-7b, with the exception of the group age fifteen through forty-four, which uses slightly less than its population share of inpatient resources, though admissions for this group exceed its population share. This suggests that the average cost per case for young adults is low in relation to the mean across all age groups. This conclusion is supported by figures 2-9a and 2-9b, which depict mean expenditures per admission and per patient-day for each age group in relation to their respective overall means. The results confirm what would be expected: older adults and the elderly are the most expensive to treat in the hospital. The implication is that as a country's population ages, if there is no significant change in the delivery of health care services, hospital costs will rise.

Diseases and Conditions Treated in Hospitals

Data on hospital admissions from six countries are summarized in table 2-5 by the primary headings of the International Classification of Diseases (ICD). Conditions relating to pregnancy and childbirth were the leading cause of admissions in five of the six countries. This is as expected. A country's policy or the population's preferences regarding hospital deliveries, however, are also important factors. Trauma care related to accidents, injuries, and poisonings was among the top five causes of admission in each country, reflecting both the relatively frequent nature of serious accidents across countries at various levels of development and the important role of hospitals in providing trauma care. Other leading causes of admission are infectious and parasitic diseases, respiratory diseases, and diseases of the digestive and genitourinary systems.

The causes of admission to hospitals can be reorganized into groups to assess the broad epidemiological composition of inpatient use. Table 2-6 presents such a reorganization into five inclusive groups. The table provides a snapshot of several countries at different stages with respect to the epidemiological transition. The relative sizes of the communicable disease and the chronic and noncommunicable disease categories are of particular interest. The relative share of communicable diseases is greatest in those countries (Nigeria and Malawi) with the highest infant mortality rates and the shortest average life span (see table A2-1). Conversely, the countries with the lowest IMR and longest average life span (Jamaica and Korea) have the greatest share of chronic and noncommunicable admissions. These statistics are not simply a function of a country's level of development, as measured by per capita GNP; Oman has the highest per capita GNP, followed by Korea, Belize, Jamaica, Nigeria, and Malawi. Communicable diseases are, in general, more amenable to broad-based primary prevention efforts than are noncom-

Figure 2-9a. Cost per Admission in Relation to the Population Mean, Korea, 1986



Source: FKMIS 1987.

Figure 2-9b. Cost per Patient-Day in Relation to the Population Mean, Korea, 1986

Source: FKMIS 1987.

Table 2-5. Ten Leading Causes of Admission to Hospitals, Selected Countries

Rank	Belize (1985)[a]	Jamaica (1985)[b]	Korea (1986)[c]	Malawi (1986)[d]	Nigeria (1984)	Oman (1984)[e]
1	Pregnancy/ childbirth (40.8%)	Pregnancy/ childbirth (42.9%)	Pregnancy/ childbirth (25.9%)	Pregnancy/ childbirth (29.8%)	Parasitic/ infectious (31.3%)	Pregnancy/ childbirth (30.2%)
2	Accidents/injuries (9.8%)	Accidents/injuries (11.4%)	Digestive (15.3%)	Parasitic/ infectious (26.9%)	Pregnancy/ childbirth (23.1%)	Parasitic/ infectious (18.7%)
3	Parasitic/infectious (8.6%)	Digestive (6.0%)	Respiratory (9.8%)	Respiratory (9.1%)	Respiratory (9.8%)	Respiratory (13.0%)
4	Respiratory (6.8%)	Genitourinary (5.9%)	Parasitic/infectious (7.0%)	Anemias (6.6%)	Genitourinary (5.8%)	Accidents/injuries (8.1%)
5	Digestive (6.2%)	Heart/circulatory (5.9%)	Accidents/injuries (6.9%)	Accidents/injuries (5.0%)	Accidents/injuries (5.3%)	Genitourinary (5.8%)
6	Heart/circulatory (3.7%)	Respiratory (5.2%)	Neoplasms (6.6%)	Endocrine/ metabolic (4.1%)	Digestive (5.0%)	Senility/ill-defined (4.7%)
7	Perinatal (3.0%)	Parasitic/ infectious (4.8%)	Genitourinary (5.8%)	Genitourinary (3.3%)	Nervous/sensory (3.3%)	Digestive (4.6%)
8	Genitourinary (2.3%)	Neoplasms (3.3%)	Heart/circulatory (5.8%)	Skin (2.9%)	Anemias (3.0%)	Heart/circulatory (4.1%)
9	Endocrine/ metabolic (2.1%)	Endocrine/ metabolic (2.5%)	Nervous/sensory (3.2%)	Nervous/sensory (2.3%)	Endocrine/ metabolic (2.8%)	Nervous/sensory (2.7%)
10	Skin (2.0%)	Skin (1.6%)	Perinatal (2.7%)	Digestive (2.2%)	Skin (2.4%)	Skin (1.8%)

a. All public hospitals.

b. All public acute care hospitals except 200-bed children's hospital and 500-bed university hospital.

c. Data on those covered by Industrial Employment Medical Insurance, Medical Insurance for Government Employees and Private School Teachers, and Compulsory Regional and Occupational Medical Insurance (99 percent of the insured and 46 percent of the total population in 1986).

d. All public hospital inpatients.

e. All MOH facilities.

Source: Raymond and others 1987 (Belize); GOJMOH 1987 (Jamaica); FKMIS 1987 and KMIC 1987 (Korea); Mills 1989 (Malawi); World Bank data (Nigeria and Oman).

Table 2-6. Distribution of Causes of Hospital Admissions across Major Categories of Conditions, Selected Countries
(percent)

Country	*Year*	*Pregnancy and perinatal*	*Communicable diseases*[a]	*Chronic and non-Communicable diseases*[b]	*Accidents, injuries*	*Other c*
Belize	1985	44	15	22	10	9
Jamaica	1985	44	10	30	11	5
Korea	1986	29	17	45	7	3
Malawi	1986	31	36	26	5	3
Nigeria	1984	23	41	22	5	8
Oman	1984	31	32	24	8	5

a. Infectious and parasitic diseases and respiratory diseases.

b. Noncommunicable diseases include neoplasms; endocrine, nutritional, and metabolic diseases; anemias; mental disorders; and diseases of the nervous system and sense organs, circulatory system, digestive system, the skin and musculoskeletal system, and the genitourinary system.

c. Congenital anomalies and ill-defined conditions.

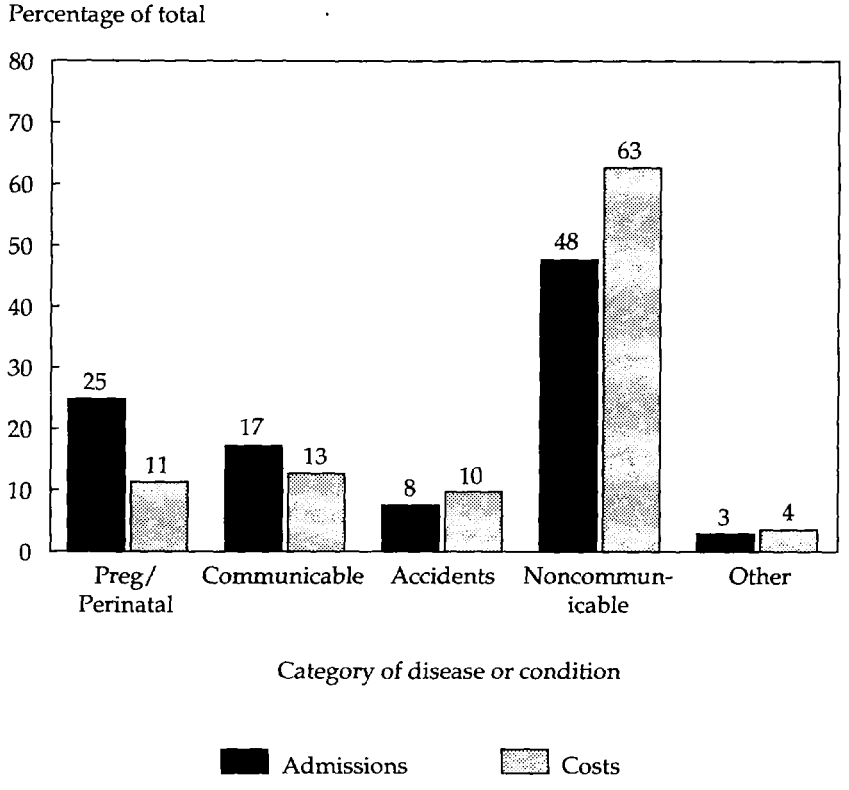
Source: Raymond and others 1987 (Belize); GOJMOH 1987 (Jamaica); FKMIS 1987 and KMIC 1987 (Korea); Mills 1989 (Malawi); World Bank data (Nigeria and Oman).

municable diseases, and thus the relative share of these diseases in total hospital admissions in part reflects the effectiveness of a country's primary health care (and specifically, communicable disease control) policies, in addition to a country's underlying demographic and epidemiological characteristics.

Countries that have been successful in controlling the spread of communicable diseases and that have experienced increases in the average length of life of their citizens are faced with a new problem: a greater percentage of persons with chronic and noncommunicable diseases than existed previously. In figures 2-8, 2-9a, and 2-9b we showed that, in Korea, the cost of treating a person over age forty-five was expensive compared with the cost of treating a younger person. In figure 2-10 we again use Korean insurance data to compare the percentage of total admissions by category of condition with the percentage of total costs for the same categories.[7] The figure indicates that communicable diseases and pregnancy and perinatal conditions are relatively inexpensive admissions, whereas admissions for accidents and noncommunicable diseases are relatively expensive. The cost data suggest that more than 70 percent of inpatient care resources are used for patients in these latter two categories.

The distribution of admissions can be used to portray relative resource use across disease categories if the percentages are adjusted for case mix

Figure 2-10. Distribution of Admissions and Inpatient Costs by Category of Principal Condition, Korea, 1986



Source: KMIC 1987.

(that is, resource use per case of a particular type). The Korean data on admissions and cost by disease categories were used to develop case mix weights. These weights were applied to the distribution of admissions by disease in selected countries to approximate relative resource use by condition. The estimates are only indicative because they apply results from one country to other countries. The weights were generated by dividing the percentage of total costs for a disease category by the percentage of total admissions for that same category. The percentages of admissions for a country were then multiplied by the appropriate weight for a disease or condition to generate a distribution of inpatient resource use (corrected to sum to 100 percent). The results of this exercise for six countries are presented in table 2-7.

Although the results in table 2-7 are only indicative, they suggest two important conclusions: (1) in all countries, a greater share of hospital resources is devoted to treating patients with noncommunicable conditions and for trauma care than is indicated by data on total admissions, and (2) as the demographic and epidemiological transitions proceed, the demand on hospital resources will increase as the mix of patients becomes more expensive to treat.

The implications of the epidemiological transition for the health systems of Mexico, Brazil, and China have recently been analyzed (Frenk and others 1989; Briscoe 1990; and Bumgarner 1992, respectively), and the findings are consistent with the conclusions reached above. In the Mexico study, Frenk and his colleagues call for a transition in the health system that involves proactive, innovative primary health care measures to meet the changing pattern of service needs. In the Brazil study, Briscoe warns that the demand for individual treatment of "post-transition" diseases will escalate and may divert resources from preventive efforts aimed at these as well as at communicable diseases, which will remain

Table 2-7. Distribution of Hospital Costs across Major Categories of Conditions, Based on Case Mix–Adjusted Admissions, Selected Countries (percent)

Country	*Year*	*Pregnancy and perinatal*	*Communicable diseases*[a]	*Chronic and noncommunicable diseases*[b]	*Accidents, injuries*	*Other*[c]
Belize	1985	24	14	34	15	13
Jamaica	1985	23	9	44	17	7
Korea	1986	13	13	61	9	4
Malawi	1986	17	32	39	8	4
Nigeria	1984	12	35	34	8	11
Oman	1984	16	27	37	12	8
Case mix weights		0.46	0.74	1.31	1.30	1.25

Note: The percentage distribution of admissions for each country was first multiplied by the appropriate case mix weight. The resulting percentages were then normalized to sum to 100 percent for each country.

a. Infectious and parasitic diseases and respiratory diseases.

b. Noncommunicable diseases include neoplasms; endocrine, nutritional, and metabolic diseases; anemias; mental disorders; and diseases of the nervous system and sense organs, circulatory system, digestive system, the skin and musculoskeletal system, and the genitourinary system.

c. Congenital anomalies and ill-defined conditions.

Source: Raymond and others 1987 (Belize); GOJMOH 1987 (Jamaica); FKMIS 1987 and KMIC 1987 (Korea); Mills 1989 (Malawi); World Bank data (Nigeria and Oman).

important in many parts of the country. In his study of the health system in China, Bumgarner conservatively estimated that, solely because of the epidemiological transition, the annual rate of growth of per capita health care costs will be 2 percent higher than that of per capita GNP. These studies emphasize the need to allocate resources to chronic disease prevention and control programs immediately as a means to avoid future preventable loss of life and keep hospital treatment costs from reaching completely unmanageable levels.

Cost-Effectiveness of Hospital Services

This section contains a brief review of the relative cost-effectiveness of hospital and nonhospital health care services. Both hospitals and PHC programs cover multiple and often overlapping activities, and it is difficult to assess the effectiveness of all interventions collectively. "Primary health care" has become a particularly amorphous term, pervasive in the literature but difficult to define in operational terms. The term is nevertheless useful because its connotation of community-level delivery programs contrasts with the connotation of hospital services delivered through large facilities socially detached from the community. Primary health care also connotes prevention rather than cure, although the term does include simple curative care such as oral rehydration for diarrhea and first-level curative contact for other health problems. In chapter 6 we argue that hospitals and PHC should be more integrated, and the services provided by the health sector should be balanced and interlinked, from lower-level preventive and curative outreach programs to upper-level facilities. The question remains, however, of the appropriate balance of services within the integrated system. In order to provide one dimension of an answer to this question, we discuss below the individual activities carried out by hospitals and nonhospital programs.

It is difficult to discuss the cost-effectiveness of health services without a knowledge of the epidemiological and resource environment because the relative effectiveness of services changes with the context. The discussion below distinguishes loosely between low-resource countries with comparatively high mortality rates (say, lower-middle-income countries and below that have infant mortality rates above fifty) and high-resource countries with low mortality rates. Our interest encompasses mortality and morbidity for all age groups, but health status has been found to correlate roughly with infant mortality. There are a few exceptions, but generally, high-mortality countries have fewer health resources, and low-mortality countries have relatively greater resources. The exceptions—countries that have fewer resources but have nevertheless achieved lower levels of infant mortality—are instructive. For example, China, Sri Lanka, and Costa Rica are low- and lower-middle-income

countries that have achieved relatively low infant mortality rates through innovative and encompassing PHC programs plus investments in education, nutrition, clean water, and sanitation (Halstead, Walsh, and Warren 1985).

If health planners working in poorer countries were asked to rank health interventions by their efficiency in achieving decreased morbidity and mortality, they would almost universally place primary health care, especially basic services delivered through outreach or rural health centers or health posts, near the top and large, urban-based institutional facilities near the bottom, with regard to both cost-effectiveness and equity. This consensus view derives from the fact that health problems targeted by primary health care programs are epidemiologically the most important, especially to low-income groups, in low- and middle-income countries.

Table 2-8 contrasts the epidemiological pattern of diseases in Ghana, China, Mexico, and the United States as examples of high- and low-mortality countries. Ghana (1979) provides an example of an epidemiological environment typical of many low-income and high-mortality countries.

Table 2-8. Total Days of Life Lost by Major Category of Disease in Selected Countries
(percent)

Category	*Ghana (1979)*	*China (1985)*	*Mexico (1985)*	*United States (1988)*
Infections, respiratory and digestive diseases, malnutrition	58	25	22	—
Chronic and cardiovascular diseases, malignancies, psychiatric disorders	21	34	19[a]	39
Newborn, pregnancy, gynecological complications	15	9	14	6
Injuries, accidents, homicide, suicide	5	27	23	31
Other	1	5	22	24[b]
Total	100	100	100	100

— Not available.

a. Malignancies and cardiovascular only.

b. Includes infectious diseases, which represent approximately 5 to 10 percent of total days lost.

Source: GHAT 1981 (Ghana); People's Republic of China 1986 (China); Hijar-Medina 1990 and Cavazos-Ortega and others 1989 (Mexico); Centers for Disease Control 1989 (United States).

The leading causes of morbidity in Ghana and other high-mortality, low-resource countries are upper respiratory illness, diarrhea, parasitic diseases, and accidents. The leading causes of mortality are vaccine-preventable diseases, respiratory diseases, malnutrition, diarrhea, and accidents. With the exception of accidents, hospitals do not play a dominant role in reducing lost years of life from these causes. Accumulating studies (see, for example, Barnum, Tarantola, and Setiady 1980; Feachem 1986; and Shepard, Brenzel, and Nemeth 1986) demonstrate that preventive measures, such as immunization and prenatal care, and simple curative measures, such as oral rehydration, can be delivered efficiently through rural health post and outreach programs and are only a fraction of the cost of the alternative inpatient care where it is available.

The epidemiological picture in China in 1985 illustrates the pattern of diseases in a low-mortality country. After earlier success at reducing mortality from infectious diseases and lowering fertility, chronic diseases have emerged as a significant health problem. Accidents and injuries have also become important sources of morbidity and mortality. Increasingly, with a lowering of overall mortality, the pattern of diseases can be expected to resemble that in the industrial countries (the 1988 pattern in the United States of years of life lost is given for reference in the last column). Hospitals play a somewhat larger role in addressing the problems of chronic disease than in treating infectious diseases. Prevention and primary health care programs, however, still have a central role in determining the disease pattern of low-mortality countries.

Hospitals do play an essential role in the delivery of a program of coordinated health services and provide an essential backup and credibility for primary health care programs in both low- and high-mortality countries. In particular, as we will argue in more detail in chapter 6, more effort to integrate lower-level hospitals could greatly increase the effectiveness of outreach and community-based programs. Central-level hospitals can also provide technical support for lower-level services and a focus for training of skilled manpower. Nevertheless, as routinely applied, hospital services, especially in upper-level hospitals, are less cost-effective in reducing mortality or morbidity than many alternative uses of health sector resources, as shown below.

The cost-effectiveness of an array of alternative health interventions for primary and secondary prevention and treatment is summarized in tables 2-9a and 2-9b. The effects are measured, depending on the study and availability of data, by years of life gained (YLG) from prevention of premature mortality or, if the required additional morbidity data are available, by healthy years of life gained (HYLG) or, if an index of the quality of health status has been constructed, by quality-adjusted life years (QALY).

Table 2-9a. Approximate Cost-Effectiveness of Selected Primary and Secondary Prevention Activities in Health
(percent of GNP per capita)

	Low-income, high infant mortality countries[a]		*High-income, low infant mortality countries*[b]	
Intervention	*Cost per discounted YLG*	*Cost per discounted HYLG*	*Cost per discounted YLG*	*Cost per discounted QALY*
Primary prevention activities				
Immunization				
EPI package[c]	3	2		
Measles (alone)	5	5		
Polio (alone)	120	21		
Hepatitis B	20			
Cholera	16	14		
Diarrhea prevention				
Weaning education	9	8		
Breast-feeding	7	6		
Prevention of new smoking starts	2		2	

Tropical disease vector control			
Malaria	22	12	
Schistosomiasis	46	26	
Onchocerciasis		85	
Secondary prevention activities			
Prenatal screening and high-risk delivery (maternal death only)	19		
Cervical cancer screening	25		26
Breast cancer screening			
Physical exam			12
Mammography added			150
Hypertension screening			
Mild hypertension (90–110 mm Hg)			140
Moderate hypertension (110 mm Hg)			70
Hypercholesterolemia screening (drugs)			500

Note: All notes appear at the end of table 2-9b.

Table 2-9b. Approximate Cost-Effectiveness of Selected Treatment Activities in Health
(percent of GNP per capita)

Intervention	*Low-income, high infant mortality countries*[a]		*High-income, low infant mortality countries*[b]		*Estimated foreign content in lower-middle income country*[c] *(proportion)*
	Cost per discounted YLG	*Cost per discounted HYLG*	*Cost per discounted YLG*	*Cost per discounted QALY*	
Diarrhea treatment					
Oral rehydration	5	5			
Intravenous therapy	37	33			
Tuberculosis treatment					
Outpatient (rifampicin)	6	6			
Inpatient-outpatient	16	15			
Hospital treatment	40				
Neonatal intensive care:					
1,000–1,499 grams		400[d]		70	0.3
500–599 grams		3,000[d]		500	0.5
Cancer treatment					
Cervix	40		10		0.2
Breast	40		11		0.2
Colon and rectum	140		50		0.2
Lung	1,300		220		0.5
Stomach	2,700		460		0.5
Liver	4,000		660		0.5

Hip replacement	50[d]	12	0.5
Hemodialysis			
Hospital	1,000[d]	300	0.7
Home	1,000[d]	200	0.7
Coronary treatment			
Pacemaker	60[d]	12	0.7
Valve replacement	80[d]	15	0.7
Coronary bypass			
Severe angina (left vent.)	90[d]	25	0.7
Moderate angina (2 vessel)	300[d]	60	0.7

Note: The estimates should be regarded as approximate. The intention is to allow an order of magnitude comparison among broad categories of interventions. Local conditions may cause great variance in actual cost-effectiveness across countries. Cost is expressed as a percentage of GNP per capita. The purpose of using the percentage of per capita GNP rather than monetary units is to reduce program costs across countries to roughly comparable units. The measure is deficient in that it primarily adjusts for labor cost differences among countries but does not account well for differences in foreign supply costs or productivity. The deficiences are offset, however, by the convenience of the measure. Sources for the GNP and exchange rates are various years of the World Bank, *World Development Report*, and IMF, *International Financial Statistics*.

Many of the sources are reviews of cost-effectiveness rather than primary sources. If several studies gave the cost-effectiveness of an intervention, as was the case for oral rehydration therapy or immunization, an average was taken and outliers were excluded. Basic information in the literature is reported variously as cost per undiscounted or discounted years of life lost, per healthy years of life lost, or per death prevented. Several procedures were used to convert the information to the comparable measures used in the table. Undiscounted results or results reported using a different discount rate were converted to discounted units using a 3 percent discount rate. Deaths prevented were converted to years lost using life expectancies for the original time and place of the study and information on average age of death from the Ghana or China data sets summarized in table 2-8. Ratios of years of life lost from death to total healthy years lost in the Ghana study were used to convert years of life lost to healthy years in other studies.

a. IMR >50.

b. IMR <50.

c. The foreign exchange proportion is approximate. The foreign exchange requirements of hospital services are expected to vary with the size and level of development of the country. No studies have specifically examined the foreign exchange content of hospital services by function in low-income coun-

(Table continues on the following page.)

Table 2-9b (continued)

tries. In a study of hospitals in Tanzania, Dunlop (1984) estimated that 40 percent of total government hospital costs entailed foreign exchange expenditures in 1979. This is an average, however; basic services would have a smaller foreign exchange content and more technically complex services would far exceed 40 percent because of the need for special skills and training.

d. Estimated cost per discounted QALY based on results for high-income countries.

Source: Barnum, Tarantola, and Setiadi 1980; Robertson 1985; Shepard, Sanoh, and Coffi 1986 (immunization). Feachem 1986; Horton and Claquin 1983; Shepard Brenzel, and Nemeth 1986 (diarrhea). Barnum and Greenberg 1991 (smoking). Barlow and Grobar 1986 (malaria, schistosomiasis). Prost and Prescott 1984 (onchocerciasis). Herz and Measham 1987 (prenatal screening). Barnum and Greenberg 1991 (cervical cancer screening, breast cancer screening). Torrance 1986 (hy pertension screening, coronary treatment, neonatal care, dialysis, hip replacement). Williams 1985 (coronary treatment, neonatal care, dialysis, hip replacement). Barlow 1976 (general hospital treatment). Mills 1985; Mills and Drummond 1987; Drummond 1985 (additional sources).

The three measures are not strictly comparable. The HYLG measure primarily has been applied to prevention, whereas the QALY measure has been applied to treatment. For any given preventive intervention, an ordering of the magnitude of the measures will, by definition, give HYLG ≥ YLG.[8] In contrast, for a given treatment, an ordering of the magnitudes will, by definition, give, YLG ≥ QALY.[9] Thus if the cost per QALY or YLG for a first intervention is less than the cost per HYLG for a second, the cost-effectiveness of the first is evidently greater even though different measures have been used in reporting the effects. In any case, if the cost per unit measure of a first intervention is an order of magnitude greater than the cost per unit measure of a second (that is, the difference is sufficiently great that it cannot be due to differences in the technical definition of the measures), then the cost-effectiveness of the first can be accepted as a practical conclusion.

Costs are measured as a percentage of the GNP per capita (percent GNPN) at the time and in the country of the study. The purpose of using the percentage of per capita GNP rather than monetary units is to facilitate comparisons across countries. Cost as a percentage of GNP gives an intuitively clear measure of the cost of the intervention in terms of the resources used in relation to the productive capacity of the country. The measure adjusts primarily for differences in labor costs and local supply costs between countries, however, and does not account well for differences in foreign supply costs or productivity. Thus, the degree of comparability of the cost-effectiveness estimates expressed in percent GNPN is limited; differences in technical capacity and in productivity are substantial between countries, and there are important technical interventions such as complex surgery in high-income countries that can be replicated in poorer countries only at a substantially greater cost in terms of per capita GNP because of the need to use imported materials and technical training. The use of common monetary units, such as dollars, is even more problematic because differences in resulting cost estimates may reflect differences in exchange rates and wage rates more than the content of services. Conceptually, the separate components of intervention costs could be corrected for price differences across countries, but it would be difficult to do so because of the lack of appropriate indexes across the range of countries and dates in the cost-effectiveness studies surveyed.

In recognition of this problem in comparability, tables 2-9a and 2-9b are divided into two sections; the first two columns give the results of studies from low-income countries and the third and fourth columns show the results of studies carried out in high-income countries. The results in table 2-9a, for primary and secondary prevention, respectively, are not greatly affected by this distinction because most of the studies on which the table is based were carried out in developing countries, and

the foreign exchange costs for these interventions is low. For neonatal intensive care, cancer, hip replacement, hemodialysis, and heart treatments, the original studies on which table 2-9b is based were carried out in industrial countries. For these interventions an estimated cost for hospital procedures in percent GNPN for an average lower-middle-income country has been computed based on the estimated foreign exchange component in the last column and calculation of a weighted sum of the foreign and local costs.[10]

The general order of magnitude of the difference in cost-effectiveness between hospital services and primary health care programs can be established, and the dominance of lower-level interventions, especially prenatal care, diarrhea control, and immunizations is clear from the tables. Looking first at diseases that are of primary importance in high-mortality countries, we see that the cost per year of life gained from the expanded program of immunization (EPI) package or measles vaccination varies from 3 to 5 percent GNPN, and for diarrhea control using weaning, breast-feeding, or oral rehydration therapy, the cost per year of life gained varies from 5 to 9 percent GNPN.

Obstetrics and neonatal care are of particular importance because of the relatively large share of total hospital resources used for delivery in high-fertility countries. In low-resource countries, routine delivery in hospitals is not cost-effective compared with health center or attended home delivery (see, for instance, the analysis in Barnum and others 1980). Neonatal intensive care at 400 percent GNPN for births of 1,000–1,500 grams is very expensive, and for births of less than 1,000 grams the cost of wide coverage would be prohibitive. Much of the need for neonatal care could be prevented by prenatal care programs, especially with early detection of pregnancy (see the discussion in Rao 1990). A potentially cost-effective use of hospital services is for high-risk deliveries, with an estimated cost per year of life gained of about 20 percent GNPN. Few countries, however, currently have an effective screening program for high-risk deliveries or adequate transportation to provide access, and the ability to mount such a program depends crucially on the development of the supporting primary health care infrastructure.

Cesarean deliveries are not mentioned in the tables but have great implications for the misuse of hospital resources. When performed in conjunction with high-risk screening, cesarean deliveries can be an important component of programs to reduce maternal mortality. Rapid growth in the number of cesarean deliveries in some middle-income developing countries (Bobadilla and Walker 1991), and in industrial countries (OECD 1987), has raised the question of inappropriate use. In selected countries in which medical training and the financial reimbursement system provide adverse incentives, the proportion of total deliveries that are cesarean has become a potential hazard to maternal health

and an unwarranted financial burden on the health system. Brazil provides a dramatic example. The cesarean rate (1981) is 17 percent for low-income households and climbs to 58 percent for upper-income households, with an overall average of 31 percent for the country (Saxenian forthcoming). Clinically, less than 15 percent of all deliveries, on average, can benefit from a cesarean section. For routine, normal deliveries, which comprise more than 85 percent of births, cesarean sections are of greater mortality and morbidity risk for the mother than normal vaginal delivery. A World Bank country study (Saxenian forthcoming) estimates the cost of unwarranted cesareans in Brazil to be US$53 million per year.

In countries that have experienced the epidemiological transition from high to low infant and child mortality, the cost-effectiveness of hospital services increases, but not dramatically. As noted earlier, increased life expectancy greatly shifts the mix of diseases with which the health system must contend, and problems of adult ill health, especially accidents and chronic diseases, become even more important than they already are. Columns three and four of tables 2-9a and 2-9b summarize the cost-effectiveness of a number of health interventions in high-resource countries. Selected hospital services are relatively cost-effective, but preventive measures remain crucially important.

Primary prevention programs for smoking are clearly dominant among interventions for noncommunicable diseases, and hepatitis control is also a priority area compared with the application of hospital resources to lung, stomach, or liver cancers. At an estimated cost per year of life saved of 2 percent GNPN, antismoking programs are much more cost-effective than the use of hospital resources to treat the associated lung cancer or heart ailments over the lifetime of the disease. Treatment of lung cancer or heart disease costs 200 percent GNPN per year of life saved and 25 to 60 percent GNPN per QALY in a high-income country and 1,300 percent GNPN per year of life saved and 90 to 300 percent GNPN per QALY in a low- or middle-income country. Similarly, hepatitis B immunization in a high-prevalence country at 20 percent GNPN per year of life saved is more cost-effective than the 660 percent GNPN per year of life saved from treating the associated primary liver cancer.

Secondary prevention programs, such as those for cervical and breast cancer and tuberculosis, can also be cost-effective if the incidence of the targeted diseases is high and sufficient hospital infrastructure is available for follow-up. Several important secondary prevention programs (for example, prenatal care, screening for high-risk delivery, and screening for breast and cervical cancer) require the use of hospital resources for follow-up treatment of those who test positive. In fact, the cost and effects of screening and treatment are intertwined, and in some cases it is difficult to isolate screening and treatment for separate cost-effectiveness evaluation.

Cervical cancer provides a good example. The low cost per year of life saved from cervical cancer treatment is an effect of the long existence of screening programs in industrial countries. These programs have resulted in earlier detection and, if treatment is undertaken, improved survival. Thus, the estimated cost-effectiveness of treatment, which is based on improvement in survival rates, is increased for cervical cancer. As a complement, the secondary prevention program is cost-effective partly because treatment for cervical cancer involves simpler and less costly procedures when the cancer is detected early. The relation of secondary prevention and the appropriate design of a referral system is considered further in chapter 6.

Unfortunately, most studies on the cost-effectiveness of hospital services have been of technical procedures, such as dialysis, that have generated controversy in industrial countries. Most of these procedures are clearly not cost-effective in developing countries and are performed, if they are performed at all, primarily in larger provincial or central facilities. Only a few studies, such as that carried out by Barlow (1976) in Morocco, have addressed the general cost-effectiveness of relatively basic services provided in hospital facilities. Recast in percent GNPN, the cost per year of life gained for hospital services in the Morocco study was 40. This cost, although higher than many nonhospital interventions, is an average of high- and low-cost hospital activities and suggests that more detailed analysis of less technically complex hospital interventions in low-income countries could demonstrate their relative cost-effectiveness.

Marginal Cost-Effectiveness of Hospital Resources

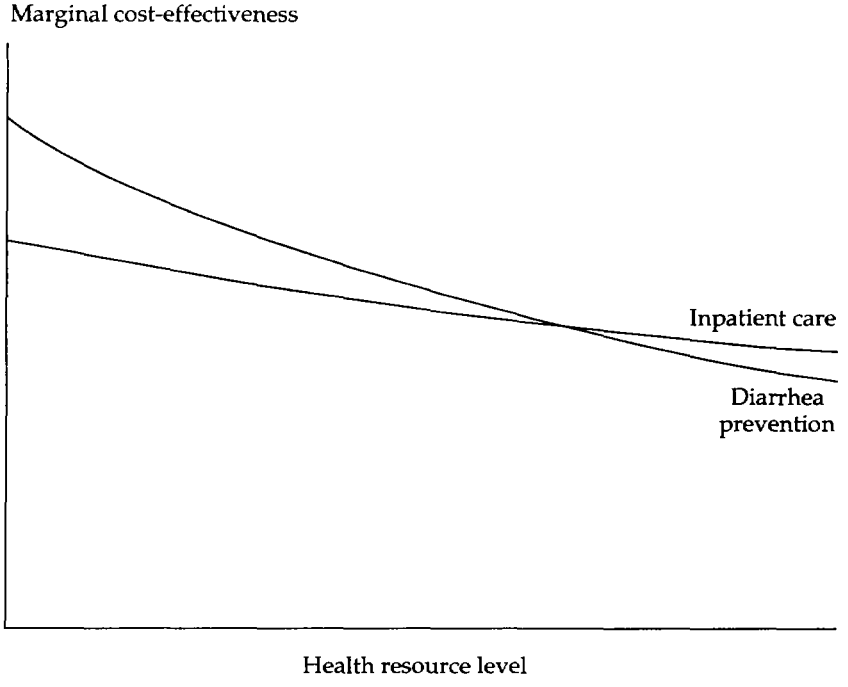
The marginal cost-effectiveness of a given activity can be defined as the change in the number of discounted healthy life years gained with a change in the expenditure on the activity. The studies cited in the discussion above do not distinguish between marginal and average cost-effectiveness. Allocational efficiency requires that the marginal cost-effectiveness of interventions be equal, and the use of average cost-effectiveness introduces bias into the comparisons. This bias can be especially important in comparing interventions in environments of low mortality and high resources, because the marginal cost-effectiveness of any intervention falls as the incidence of its related disease falls and the level of coverage by the given intervention as well as other interventions increases. In evaluating health care interventions in countries with low levels of resources and high disease rates, the difference between resource allocations based on average and marginal effects is not great. The distinction may be important, however, in resource environments in which the coverage with more basic interventions may be high.

A study of the optimum use of resources to improve child survival demonstrated that marginal cost-effectiveness can change rapidly as the coverage and use of interventions increases, with the result that the optimum intervention mix changes at alternative resource levels (Barnum and others 1980). At low-resource levels and with an IMR greater than 100, the activities with the highest marginal cost-effectiveness in improving child survival are outreach programs promoting nutrition, breast-feeding, antidiarrheal measures, prenatal care, and attended home delivery. At middle- to upper-resource levels, hospital outpatient programs and inpatient delivery become increasingly cost-effective. Only at upper levels of resources, as the IMR falls well below fifty, does inpatient care become cost-effective, and then only for selected uses. The principle remains sound and can be extended to resources used by adults as well as children and to morbidity as well as mortality. Preventive and primary curative programs should be maintained at all resource levels, but hospital inpatient care becomes increasingly cost-effective as the general mortality level falls and greater health resources become available.

If a country were following an optimum allocation strategy as the health sector developed, the marginal cost-effectiveness of any given primary health care intervention, even ones with as high an average cost-effectiveness as the diarrhea prevention activities listed in table 2-9a, would decrease with each additional unit of resources until at some point it fell below inpatient hospital care. The principle is illustrated in figure 2-11. If we follow the two lines depicting the marginal cost-effectiveness of diarrheal disease prevention and inpatient care for the case in which a country has initially invested only a limited amount of resources in hospitals, we can see how the required investment in hospitals might change as increasing resources become available. The change in the number of years of life gained with, say, an additional expenditure on diarrhea prevention or hospitals is measured on the y-axis and the level of resources used is given on the x-axis. A low level of resources used for diarrhea prevention results in great marginal cost-effectiveness compared with an equivalent expenditure on inpatients. As the resources used for diarrhea prevention activities increase, the marginal cost-effectiveness of the activities decreases until at some point (at high resource levels) diarrhea prevention becomes less cost-effective than an additional unit of inpatient care.

In most resource-poor countries, however, the health sector investment program has not followed an optimum path, and the countries have already invested heavily in hospitals even though the level of sectoral development is low. Investments in nonhospital health care programs are vastly more cost-effective in these countries. This situation is depicted in figure 2-12. As diarrhea prevention activities increase, the relative cost-effectiveness of hospitals nevertheless remains low. In these

Figure 2-11. Change in Marginal Cost-Effectiveness with Low Initial Investment in Hospital Services

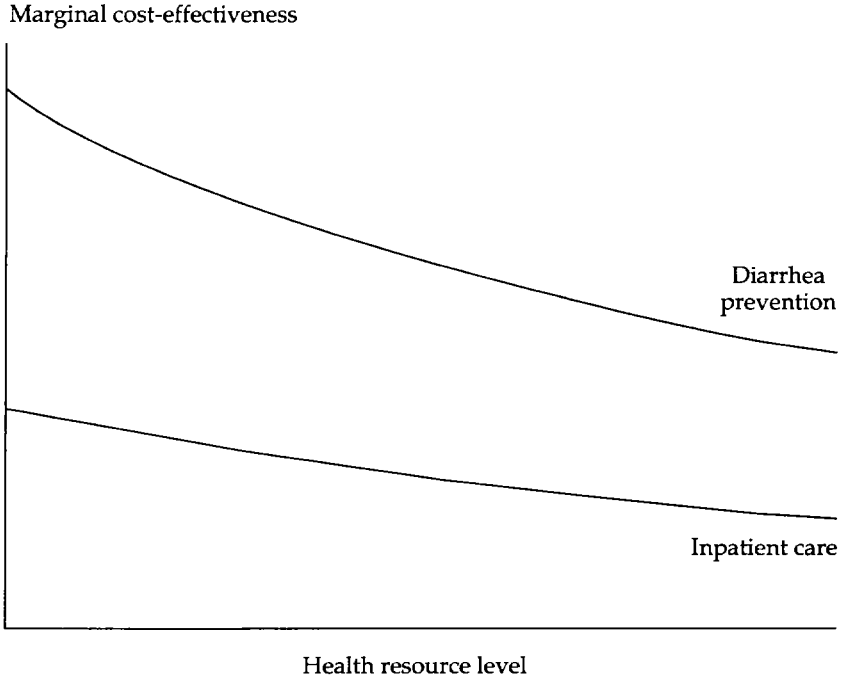


countries it becomes especially important to release public resources for more cost-effective programs by finding ways of diminishing the government's financial responsibility for hospitals. In practice it is often politically or technically difficult actually to reduce recurrent expenditures on hospitals and transfer the funds to nonhospital programs. Changes must be made at the margin by altering investment patterns; this strategy can have significant effects in time. An alternative method would be to alter the function of existing hospitals to integrate them more fully into primary care programs.

Summary

In this chapter we have surveyed the use of hospital resources in developing countries with respect to their distribution within the government health sector, their distribution within and across hospitals, their dis-

Figure 2-12. Change in Marginal Cost-Effectiveness with High Initial Investment in Hospital Services



tributional equity, and their cost-effectiveness. We have arrived at the following conclusions:

- In nearly all countries, the largest share of public sector health expenditure is for hospitals, regardless of a country's health status and income level.
- Within the hospital subsector in many countries, tertiary hospitals use a very large share of public resources in relation to district-level hospitals. The relatively limited resources available for district-level hospitals can be detrimental to service quality and impedes the function of district hospitals as the institution of choice for first referral.
- The benefits of hospital services are not distributed equitably throughout the population but instead are received disproportionately by residents of urban areas. There is also limited evidence to suggest that hospital services are used less by the poor, with the possible

exception of hospital services covered by social security in those countries in Latin America where such coverage is high.

- Nonhospital interventions are both more cost-effective and more equitable as a means of improving health status for most of the prevalent health conditions in low-income countries.

These conclusions are not surprising and are consistent with widely held views of hospital resource use. Nevertheless, the discussion leading to the conclusions is important because it marshals available evidence to sustain the prevailing consensus and underlines the importance of finding solutions to the problem of heavy hospital resource use. Within these broad and obvious conclusions the chapter reveals a number of less obvious details that have implications for health sector resource planning.

Distribution of Resources within the Health Sector

The large share of recurrent resources going to hospitals is only indicative of the actual priority placed on other health programs or sector strategies. The hospital share does not, in itself, demonstrate a country's level of support for primary health care or the effectiveness of its primary health care strategy. Although exceptionally large expenditures for hospitals are competitive with PHC programs, lower-level hospitals can provide substantial primary health care services directly as well as support to nonhospital aspects of PHC programs, as will be brought out in chapter 6. For these reasons a comparison of hospital share with health status indicators, such as the infant mortality rate, related to the goals of primary health care reveals only a modest inverse pattern with many significant exceptions. Thus, an examination of tradeoffs between hospital resource use and PHC requires country-specific analysis.

The hospital share of recurrent resources can be changed by deliberate government policies affecting either the short-term or long-term allocation. In the short term, the share may change during periods of fiscal difficulties when fewer public resources are available. Some countries have acted to protect primary care programs in the face of declining availability of recurrent resources for the entire sector, whereas others have concentrated a greater share on hospitals. Policymakers should be especially careful during periods of acute fiscal adjustment to protect the recurrent resource needs of basic health care programs.

In the long term, although the number of physicians in relation to the population appears to affect the hospital share, a more worthwhile activity than simply limiting physician supply may be to alter the medical curriculum to place a greater emphasis on primary health care and less emphasis on technically complex curative techniques. Perhaps physicians could spend more training time at district hospitals and in

primary care outreach programs. A by-product of such a change may, in the medium to long term, be a reduction in the hospital share of public sector health expenditure.

The long-term balance of recurrent resource use between hospitals and other health programs is obviously affected by past and current investments, but health sector plans often fail to project the implications of recurrent costs. Recurrent–capital cost ratios can be used to generate a rough order-of-magnitude estimate of the recurrent cost implications of current capital expenditures until project-specific financial analysis is available for more accurate projections. Whatever projection methodology is used, governments should act on the projections to coordinate government and donor investment in the sector to achieve the intended long-term balance of recurrent resource use.

Allocation of Resources within the Hospital Subsector

Tertiary-level hospitals absorb a large share of total public expenditure on hospitals. Tertiary hospitals are intended for patients with complex conditions, who are more costly to treat. It is likely that many, if not most, of the patients treated in large urban tertiary hospitals could be treated in less costly facilities but are not because the referral system is not functioning effectively and because viable alternative urban treatment centers are not available for the population living relatively close to the hospital. If concentration of funds on tertiary hospitals leads to underfunding of first-level referral hospitals, the quality of these low-level facilities is likely to deteriorate, and people who are able will tend to bypass them and seek out the nearest tertiary center. Therefore, excessive concentration of funding on a few facilities will feed on itself and result in an increasingly inefficient and inequitable allocation of hospital resources.

There appears to be a positive correlation between the share of labor in hospital recurrent expenditures and a country's level of income, though there are several exceptions. This correlation is expected on economic grounds as long as the technical possibilities for substitution between personnel and other inputs, especially drugs and supplies with relatively fixed real prices across countries, are relatively limited. Although there may be a greater relative quantity of labor used per unit of output in low-wage countries, it is not enough to cancel out the price effects of imported inputs on the relative expenditure shares. As a result, drugs and medical supplies absorb a larger share of total recurrent expenditures than in high-wage countries. This high-cost, low-productivity environment leads to a recommendation that hospitals in low-income countries should not pattern themselves after those in high-income countries; rather, emphasis needs to be placed on research to develop

productive, alternative, labor-using technologies for delivering hospital services in developing countries.

In contrast to the majority of low-income countries that have, as expected, relatively low (40 to 60 percent) personnel shares, some low-income countries have very high labor shares. Examination of the hospitals in these countries reveals the cause of this discrepancy: in the face of declining resource availability in the public sector and civil service constraints on reducing the number of employees, there are very few funds available for nonpersonnel inputs after the staff has been paid. The result is extremely low productivity, because the staff then has very few nonlabor inputs available with which to produce hospital services. Greater health sector funding and more flexibility in assigning staff to nonhospital programs within the health sector are needed to increase productivity.

Distributional Equity of Hospital Use

The use of hospital resources is important for the relative welfare of different population groups distinguished by income, geographical location, and age. Given that hospitals absorb a large share of health sector resources but can provide services to relatively few persons, these distributional implications are of great importance in setting policy. Surprisingly little information is available on this question, and there is a great need for additional research, possibly based on household surveys or comparisons of the characteristics of patients discharged from hospitals with wider demographic and income information. The limited information on hospital use by income groups indicates, but not conclusively, that bias in favor of high-income groups exists. Stronger evidence exists of geographical bias and of greater use of hospital outpatient and, to a somewhat lesser extent, inpatient services by urban populations than rural. This evidence suggests that there are distributional biases that favor high-income regions and urban populations. Hospitals are necessary as part of the overall health system, and economic logic dictates that they be located in areas with the highest population. The key equity issue raised by the geographical bias is whether the health needs of rural persons are being met by the services available to them, and whether the existing referral system provides them with access to more complex services when needed.

Consideration of the distribution of hospital services by age and disease or condition raises issues of future resource use related to the epidemiological and demographic transitions. Limited data from developing countries indicate that infants and the elderly comprise a greater share of admissions than they do of the population, whereas the five-through-fourteen age group comprises fewer admissions than its popu-

lation share. The available empirical evidence supports the hypothesis that greater resources are used per admission and patient-day for adults over age forty-five than for younger persons. This suggests that older people use more services and are more expensive to treat. An implication of the demographic transition is that as populations age, if services are organized and delivered as they are currently, hospital costs will rise.

This implication is made stronger by the fact that specific disease categories require greater hospital resources. The current leading causes of admission relate to pregnancy and communicable (infectious-parasitic and respiratory) diseases. These cases, however, have a relatively low average cost per admission, and noncommunicable conditions and accidents consume a far larger share of hospital resources than is indicated by the admissions data. It is difficult to disentangle the separate effects of the demographic and epidemiological transitions, but it is clear that as they occur their combined effects will lead to a more costly hospital case mix, further straining available resources, assuming health services continue to be financed and delivered as they are today. World Bank studies from Brazil and China stress the need to reallocate resources immediately in favor of chronic disease prevention and control programs in order to avert future high hospital costs that will arise when patients with these conditions present themselves at hospitals. This conclusion is applicable (to varying degrees) to many developing countries.

Cost-Effectiveness of Hospital Services

Understanding the relative cost-effectiveness of hospital services helps to determine the appropriate levels of hospital and nonhospital services that should be provided in an integrated referral system. Because of different resource endowments and relative prices across countries, a ranking of interventions by their relative cost-effectiveness may be different in different countries. Despite these differences, however, it is likely that the leading causes of morbidity and mortality in low-income, high-mortality countries are conditions that can be treated or prevented most cost-effectively through nonhospital interventions. In a low-mortality country, in which chronic disease is more important, hospital services become relatively more cost-effective, although primary care and prevention remain very important.

To recapitulate the conclusions of the chapter: hospital services have a role in providing referral services to support and complement nonhospital health programs, but in resource-poor countries the magnitude and diversity of hospital services to be provided should be limited. In most countries the resources going to hospitals appear to exceed the

amount required for allocational efficiency. Thus, to reduce the use by hospitals of public sector resources and to make additional resources available for nonhospital programs, it is important to find ways of increasing internal efficiency in hospitals, to find mechanisms to reduce the dependence of hospitals on public finance, and to develop low-cost alternatives to hospitals. The next four chapters address these problems.

Appendix 2A. Resource Use

The table on following pages provides indicators of health resource availability and health status for a cross-section of countries. Information is also provided for total health spending and government spending on hospitals. This latter information is available for only a limited number of countries.

Table 2A-1. Selected Health Indicators, Most Recent Estimates

Country	Per capita GNP[a] (U.S. $)	Life expectancy at birth[a]	Infant mortality (per 1,000 live births)[a]	Health inputs per 1,000 population			Health spending as a percentage of GNP	Hospitals as a percentage of public health spending[e]
				Physicians[b]	Nurses[c]	Hospital beds[d]		
Low-income countries								
Mozambique	81	49	137	0.03	0.17	1.09	4.4	36
Tanzania	120	49	112	0.04	0.18	1.38	3.2	
Ethiopia	122	48	133	0.01	0.19	0.30	3.6	49
Somalia	170	48	128	0.06	0.65	1.43	—	70
Malawi	173	47	147	0.09	0.32	1.54	2.7	81
Nepal	175	52	124	0.03	0.21	0.17	1.4	25
Lao P.D.R.	175	49	105	0.73	1.88	0.93	—	—
Bangladesh	182	51	106	0.15	0.11	0.28	1.7	61
Guinea-Bissau	183	40	147	0.14	0.89	1.86	—	—
Chad	188	46	127	0.03	0.29	1.31	—	—
Bhutan	195	48	125	0.10	0.33	0.71	2.0	—
Sierra Leone	209	42	149	0.07	0.92	1.21	3.0	—
Burundi	218	49	70	0.05	0.23	0.68	—	66
Madagascar	220	51	117	0.10	0.58	2.50	9.2	—
The Gambia	236	44	138	0.09	0.46	1.67	—	45
Zaire	244	53	94	0.08	0.56	1.65	—	—
Nigeria	249	51	100	0.16	1.18	0.73	7.8	—
Uganda	250	49	99	0.05	0.49	1.52	2.1	43
Mali	260	48	167	0.04	0.74	0.71	0.8	—
Niger	292	45	130	0.03	2.18	0.50	—	30
Burkina Faso	313	48	135	0.02	0.59	0.57	—	—
Rwanda	320	49	118	0.01	0.23	1.67	3.5	—

India	347	59	95	0.40	0.59	0.77	4.3	71
China	356	70	30	0.99	0.71	1.98	4.0	61
Haiti	359	55	94	0.14	0.44	0.72	—	—
Equatorial Guinea	359	46	122	—	0.80	—	—	—
Kenya	364	59	68	0.10	1.55	1.65	2.3	73
Pakistan	365	55	106	0.34	0.20	0.59	3.5	52
Sao Tome and Principe	373	66	71	0.50	3.55	—	—	—
Central African Republic	379	51	100	0.04	0.45	1.55	—	—
Benin	382	51	112	0.06	0.57	1.13	4.1	—
Ghana	383	55	86	0.05	0.60	1.57	2.4	—
Togo	394	54	90	0.11	0.90	1.43	—	—
Guyana	407	64	53	0.16	1.13	3.33	4.4	—
Zambia	412	54	76	0.14	1.34	3.53	5.6	—
Maldives	417	61	73	0.07	1.63	—	—	—
Sudan	428	50	104	0.10	0.79	0.88	6.0	—
Sri Lanka	433	71	20	0.18	0.78	2.94	2.3	70
Lesotho	437	56	96	0.05	0.26	1.67	2.0	74
Guinea	442	43	140	0.02	0.19	1.53	—	—
Comoros	456	55	94	0.08	0.45	2.10	—	—
Indonesia	503	61	64	0.11	0.79	0.55	2.4	37
Mauritania	505	46	123	0.08	0.85	0.78	—	—
Solomon Islands	577	64	49	0.13	1.69	5.68	—	—
Low-income countries with missing GNP data								
Afghanistan	—	42	170	0.16	0.11	0.27	—	—
Cambodia	—	50	121	—	0.73	1.08	—	—
Liberia	—	54	137	0.11	0.73	1.67	—	—
Myanmar	—	61	66	0.27	1.18	0.85	3.2	33
Viet Nam	—	66	43	1.06	1.68	3.70	—	—

(Table continues on the following page.)

Table 2A-1 (continued)

Country	Per capita GNP[a] (U.S. $)	Life expectancy at birth[a]	Infant mortality (per 1,000 live births)[a]	Health inputs per 1,000 population			Health spending as a percentage of GNP	Hospitals as a percentage of public health spending[e]
				Physicians[b]	Nurses[c]	Hospital beds[d]		
Middle-income countries								
Angola	604	45	132	0.06	0.99	2.72	—	—
Bolivia	628	54	106	0.65	0.44	2.00	2.3	—
Arab Republic of Egypt	639	60	68	1.30	1.28	2.09	—	—
Republic of Yemen	650	48	125	0.15	0.58	0.45	6.0	
Senegal	654	48	82	0.08	0.49	1.25	3.1	50
Zimbabwe	654	64	45	0.14	1.39	2.01	4.2	54
Kiribati	702	55	59	0.51	4.40	4.80	—	—
Philippines	714	64	42	0.15	0.37	1.70	2.4	71
Western Samoa	722	66	48	0.28	2.45	4.36	—	—
Cape Verde	778	66	41	0.19	1.40	2.19	—	—
Dominican Republic	781	67	61	0.57	0.83	2.50	—	—
Côte d'Ivoire	788	53	92	—	0.49	1.16	5.4	46
Morocco	878	61	69	0.21	0.95	1.24	1.2	—
Papua New Guinea	892	54	59	0.16	1.14	4.81	3.8	45
Swaziland	896	56	114	0.05	0.95	3.37	5.5	52
Tonga	907	67	24	0.60	1.83	3.56	—	—
Honduras	909	65	66	0.66	1.49	1.25	—	—
Guatemala	916	63	55	0.46	1.17	1.67	—	—
People's Republic of the Congo	943	54	115	0.12	1.73	4.59	9.4	—
Syrian Arab Republic	963	66	44	0.77	1.12	1.13	—	—
Vanuatu	981	64	71	0.19	2.18	6.08	—	—
Cameroon	995	57	90	—	0.51	2.50	7.1	—

Ecuador	1,023	66	61	1.22	1.64	2.50	3.4	—
Namibia	1,027	57	101	—	—	—	4.4	—
Paraguay	1,032	67	32	0.69	1.20	1.67	—	—
Peru	1,061	62	79	0.96	0.99	1.69	—	—
El Salvador	1,063	63	55	0.35	1.75	2.00	—	62
Colombia	1,212	69	38	0.80	1.52	1.65	4.9	67
Thailand	1,224	66	28	0.16	1.42	1.54	3.8	58
Tunisia	1,261	66	46	0.46	2.73	2.13	4.5	70
Turkey	1,368	66	61	0.73	0.98	2.08	3.0	63
Jamaica	1,400	73	16	0.49	2.50	3.33	4.6	72
Botswana	1,603	67	39	0.14	1.43	2.37	3.3	49
Jordan	1,640	67	52	0.90	0.79	0.94	6.8	75
Fiji	1,675	67	19	0.49	2.36	2.75	3.8	—
Belize	1,720	68	46	0.45	2.17	3.32	—	—
Panama	1,764	72	22	1.00	2.57	3.33	5.6	—
Dominica	1,764	75	17	0.32	1.88	4.44	5.3	—
Chile	1,771	72	19	0.81	2.70	3.41	6.1	—
Costa Rica	1,773	75	17	1.04	2.22	3.41	—	—
Poland	1,774	71	16	2.05	5.35	7.64	3.7	—
St. Lucia	1,810	71	20	0.26	1.90	5.07	5.0	—
Mongolia	1,824	62	64	9.81	4.72	11.35	—	—
Grenada	1,891	69	32	0.47	—	5.95	7.0	—
Mauritius	2,017	70	21	0.53	1.71	3.33	9.0	—
Mexico	2,079	69	40	0.80	1.14	1.25	3.4	58
Argentina	2,103	71	30	2.68	1.19	5.59	7.1	—
Malaysia	2,136	70	22	0.52	0.99	2.50	3.5	59
Algeria	2,287	65	69	0.43	3.30	2.63	5.4	—
Uruguay	2,448	73	22	1.95	5.29	3.25	6.4	—

(Table continues on the following page.)

Table 2A-1 (continued)

Country	Per capita GNP[a] (U.S. $)	Life expectancy at birth[a]	Infant mortality (per 1,000 live births)[a]	Health inputs per 1,000 population			Health spending as a percentage of GNP	Hospitals as a percentage of public health spending[e]
				Physicians[b]	Nurses[c]	Hospital beds[d]		
Venezuela	2,449	70	35	1.43	2.71	2.60		
South Africa	2,470	61	68	—	2.43	—		
Brazil	2,496	66	59	0.93	0.83	5.00	5.6	68
Hungary	2,585	71	16	3.26	5.76	9.17	5.4	—
Bulgaria	2,662	72	14	3.63	6.44	11.14	—	—
Yugoslavia	2,937	72	24	1.82	3.93	5.98	—	—
Gabon	2,992	53	98	0.36	3.67	1.25	—	—
Suriname	3,006	67	40	0.79	3.62	8.90	—	—
Islamic Republic of Iran	3,007	63	90	0.34	0.87	1.52	—	—
Trinidad and Tobago	3,338	71	15	1.05	3.98	5.00	—	—
Czechoslovakia	3,453	71	12	3.60	6.91	12.45	—	—
Seychelles	4,226	70	18	0.46	5.03	4.98	—	—
Portugal	4,249	75	14	2.42	1.59	5.00	6.4	44
Republic of Korea	4,400	70	23	0.87	1.72	1.68	5.1	33
Oman	5,217	65	36	0.91	2.55	1.81	3.4	—
Libya	5,308	62	77	1.44	2.85	4.83	—	—
Greece	5,346	77	12	2.85	2.24	6.16	5.3	33
Malta	5,832	73	10	1.14	8.84	10.00	—	—
Countries believed to be middle income, but with missing GNP data								
Antigua and Barbuda	—	74	20	—	3.07	6.53	—	—
Djibouti	—	48	117	0.24	1.98	3.60	—	—

French Guiana	—	73	—	1.40	16.65	12.15	—	—
Gibraltar	—	—	5	—	—	—	—	—
Iraq	—	63	67	0.55	0.58	1.85	—	—
Lebanon	—	65	47	1.49	1.59	4.34	—	—
Macao	—	72	11	1.80	2.35	5.49	—	—
Martinique	—	76	10	1.42	—	10.06	—	—
Montserrat	—	71	30	0.42	—	—	—	—
New Caledonia	—	69	33	0.66	2.41	—	—	—
Nicaragua	—	64	57	0.67	1.87	2.50	—	—
Pacific Islands Trust Territories	—	72	20	—	—	—		
Reunion	—	72	13	—	3.62	8.66	—	—
Romania	—	71	24	1.76	3.65	8.77	3.0	53
St. Kitts and Nevis	—	69	38	0.46	7.94	8.65	—	—
St. Vincent and the Grenadines	—	70	25	0.23	1.48	4.99	—	—
High-income countries								
Saudi Arabia	6,020	64	67	1.35	2.90	1.46	—	—
Barbados	6,346	75	13	0.89	4.49	8.53	—	—
Bahrain	6,380	69	33	1.22	2.71	3.31	—	—
Cyprus	7,040	76	11	1.34	3.67	5.55	—	—
Ireland	8,713	74	8	1.47	7.16	9.71	7.5	73
Spain	9,329	77	8	3.17	3.87	5.20	6.0	48
Israel	9,787	76	11	2.90	9.37	5.00	—	—
Singapore	10,352	74	7	0.76	2.96	4.05	—	—
Hong Kong	10,372	77	7	0.93	4.15	4.89	—	—
The Bahamas	11,119	68	24	0.94	4.85	4.40	—	—
New Zealand	12,067	75	10	1.74	12.39	10.13	6.8	66
Australia	14,357	77	8	2.29	8.73	12.00	7.6	56

(Table continues on the following page.)

Table 2A-1 (continued)

Country	Per capita GNP[a] (U.S. $)	Life expectancy at birth[a]	Infant mortality (per 1,000 live births)[a]	Health inputs per 1,000 population			Health spending as a percentage of GNP	Hospitals as a percentage of public health spending[e]
				Physicians[b]	Nurses[c]	Hospital beds[d]		
United Kingdom	14,612	76	9	1.64	8.34	9.33	6.0	57
Italy	15,118	77	9	4.28	4.00	10.59	7.3	52
Qatar	15,833	70	29	1.74	4.48	2.54	—	—
Netherlands	15,923	77	7	2.22	5.96	12.53	8.4	65
Kuwait	16,153	74	15	1.57	4.94	4.14	—	—
Belgium	16,223	76	9	3.02	9.26	9.38	7.3	23
Austria	17,303	76	9	2.57	5.43	11.14	8.4	18
France	17,821	77	7	3.13	9.51	7.22	8.6	55
United Arab Emirates	18,414	71	24	0.98	2.56	2.96	—	—
Canada	19,032	77	7	1.96	8.25	9.88	8.6	59
Iceland	20,411	78	6	2.30	11.49	—	7.9	70
Germany	20,442	75	8	2.65	4.42	11.50	8.1	43
Denmark	20,453	75	8	2.51	16.43	9.68	5.9	76
United States	20,910	76	10	2.12	13.58	5.85	11.1	59
Sweden	21,574	77	6	2.59	1.48	14.81	9.1	75
Finland	22,121	75	6	2.26	16.92	15.54	7.5	51
Norway	22,294	77	8	2.22	17.50	15.00	7.4	74
Japan	23,811	79	5	1.51	5.43	11.58	6.9	42
Luxembourg	26,217	75	9	1.81	4.14	11.84	7.2	29
Switzerland	29,883	78	6	1.44	7.74	11.16	7.7	56
Countries believed to be high income, but with missing GNP data								
American Samoa	—	—	—	0.83	0.51	—	—	—

Bermuda	—	—	—	0.76	—	9.31
Brunei Darussalam	—	75	9	0.55	3.78	3.11
Channel Islands	—	77	8	1.52	—	10.60
Faeroe Islands	—	74	23	—	—	8.13
French Polynesia	—	72	21	1.25	2.35	—
Greenland	—	63	29	1.14	—	14.35
Guadeloupe	—	74	15	1.39	1.90	11.14
Guam	—	73	11	1.18	4.75	—
Isle of Man	—	—	16	—	—	—
Netherlands Antilles	—	77	13	—	—	9.12
Puerto Rico	—	75	13	—	—	—
Virgin Islands (U.S.)	—	74	18	—	—	—
Other economies						
Albania	—	72	26	—	5.25	7.10
Cuba	—	76	12	1.89	3.51	4.62
Democratic People's Republic of Korea	—	70	27	2.38	—	—
Former U.S.S.R.	—	70	24	3.70	5.90	10.97

— Not available.
a. Data are for 1989.
b. Most estimates are for 1985 or later.
c. From about 1985.
d. From about 1980.
e. For OECD countries, measured as the share of public inpatient expenditures in total public sector health expenditure.
Source: OECD 1990 and World Bank data.

Notes

1. The relationship of the number of beds per capita (*Beds/Pop*) to GNP per capita (GNPN) using the data in figure 2-1 is

$$Log\ (Beds/Pop) = -9.416 + 0.477\ [Log\ (GNPN)]$$
$$(-35.47)\quad (13.45)$$

$$R^2 = 0.57 \qquad n = 139$$

(*t*-statistics are given in parentheses).

2. The log-linear relationship explaining the hospital share (*HOSP%EX*) is

$$Log\ (HOSP\%EX) = 1.230 - 0.068\ [Log(GNPN)] + 0.159\ [Log(MD/Pop)\ ]$$
$$(1.23)\quad (-0.76) \qquad\qquad (2.71)$$

$$R^2 = 0.29 \qquad n = 29\ (developing\ countries)$$

(*t*-statistics are given in parentheses.)

3. A pyramidal conceptualization of a health system is discussed in chapter 6.

4. The relationship of the personnel share in total hospital cost to GNP per capita (using the data from table 2-4), controlling for the level of hospital ($D = 1$ for district-level hospitals, 0 for other levels) is

$$Personnel\ Cost\ Share\ (\%) = 36.5 + 0.020\ GNPN + 7.30\ D$$
$$(7.31)\quad (3.48) \qquad (1.57)$$

$$R^2 = 0.33 \qquad n = 31$$

(*t*-statistics are given in parentheses).

5. The income elasticity of hospital beds gives the percentage change in hospital beds in response to a percentage change in per capita income.

6. The Jamaican data include admissions to all acute care public hospitals. The Korean data relate to those covered by Industrial Establishment Medical Insurance, which represented 60 percent of the insured population and 32 percent of the total population in 1986.

7. These data are for insured government employees and private school teachers and staff (and all dependents), who made up 10.4 percent of the population in 1986.

8. The number of healthy years of life gained (HYLG) is the sum of years of life gained from prevention of mortality (YLG) and years gained from avoidance of disability. Polio provides an example. Using the data for Ghana provided by the Ghana Health Assessment Team (GHAT 1981), we find that a total of 3.3 HYLG per thousand people per year are gained from elimination of polio. Of this total 0.6 is gained from elimination of mortality (YLG) and the remainder, 2.7, is gained from elimination of disability. Thus, in this example, the number of YLG is substantially less than the number of HYLG.

9. A QALY is, literally, a quality adjustment to a year of life gained (YLG). The adjustment yields a fraction of a year of life with full health. In the example given by Williams (1985), coronary artery bypass surgery on a patient with severe

angina and disease of the left main vessel gives a gain of 6 years compared with no surgical intervention. Adjusting these years for the loss of quality of life from the treatment and from continued chronic disability gives a gain, measured in QALYs, of 3.5 years.

10. Given an original cost expressed in U.S. dollars of $C_H^{\$}$ in a high-income country and foreign exchange costs as a proportion, f, of total costs in a lower-middle-income country, the estimated cost as a proportion of GNPN in a low-income country is:

$$C_L^{GNPN} = \frac{f \cdot C_H^{\$}}{GNPN_L^{\$}} + \frac{(1-f) \cdot C_H^{\$}}{GNPN_H^{\$}}$$

$$= \text{FOREIGN COMPONENT} + \text{LOCAL COMPONENT}$$

The superscript indicates whether the cost is measured in $ or GNPN, and the subscript indicates a high-income (*H*) or low-income (*L*) country.

3. *Hospital Costs and Efficiency*

In this chapter we provide estimates of recurrent hospital costs and consider the relation of costs to services. Our immediate objective is to provide information on the determinants of hospital costs that can be used to formulate policy recommendations concerning efficiency and financing. Related benefits that can arise from an understanding of costs include estimates of the potential savings from improved referral patterns, the relative cost-effectiveness of alternative programs, and projections of future recurrent resource requirements of current and proposed facilities.

The methodology and findings of studies that use two very different methods of cost analysis are reviewed in this chapter. The first method makes use of accounting information and reanalysis of hospital service records to examine hospital costs and performance. The second one makes use of statistical procedures to infer the relation of hospital costs to services provided. The *accounting* method can be applied usefully to a single hospital and can involve a labor-intensive, detailed examination of hospital accounts, staffing patterns, and admissions. It is also possible, although somewhat less accurate, to derive hospital accounting costs by using aggregate government budgets or expenditure data. Less detailed data are needed in the *statistical* method, but it requires observations of costs and service use for many hospitals. Statistical studies provide insights into cost issues—the relation between marginal and average cost, and the degree to which hospitals exhibit economies of scale and scope—that accounting studies do not reveal as readily. Ideally, the information used for the statistical analyses would be derived from a large number of detailed and well-documented observations. In actuality this is not often possible, and the lesser quality of data in a statistical analysis must be compensated for by inferring a general pattern of costs from a large number of observations. Thus, the accounting and statistical methods yield different but complementary views of costs.

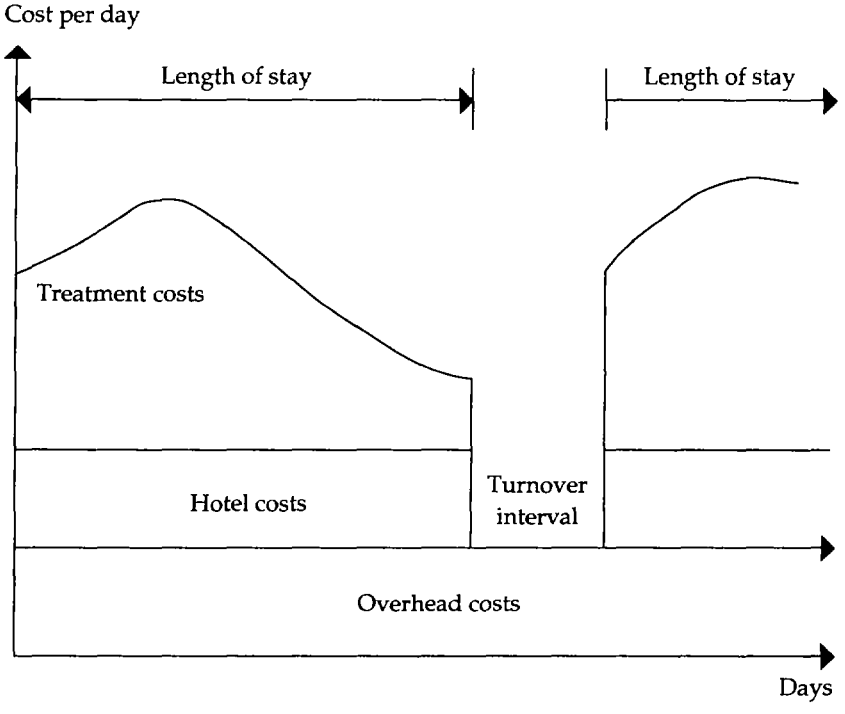
Whichever method of costing is used, the unit of analysis (such as patient-days, admissions, or outpatient visits) and consideration of both average and marginal costs are important, because the use of any single output or cost measure may produce misleading results. The reasons for this can be seen in figure 3-1, which provides a schematic description of the nature of inpatient hospital costs. As depicted in the figure, recurrent inpatient costs are considered to have three components:

- *Overhead costs.* These costs remain essentially constant regardless of whether a bed is occupied. Typically, they include items such as heating and maintenance, but for many public hospitals in developing countries, personnel may be a large component of overhead costs, because it may not be possible for staff to be reduced in the short run during periods of low occupancy. The magnitude of overhead costs is related to hospital size.
- *"Hotel" costs.* These are costs, such as catering, laundry, and linen, that are incurred for each patient-day in the hospital. They tend to be constant for each day of a patient stay, though there will be some variation related to diagnosis and patient characteristics.
- *Treatment costs.* These are case-dependent costs associated with the particular diagnostic, therapeutic, and other treatment services provided to the patient. As depicted in figure 3-1, these costs tend to peak in the first few days of a patient stay, when, for example, there might be an operation, and then diminish thereafter. The actual pattern of treatment costs in any hospital will vary depending on the clinical management of inpatients.

Figure 3-1 illustrates the importance of using various units of analysis when interpreting cost data. The figure posits that marginal costs per day equal average costs per day for overhead costs and per patient-day for hotel costs but differ for treatment costs during the period of the patient stay. If the case were the only unit of analysis, the changing nature of marginal treatment costs during a stay would not be observed. Cost comparisons between two hospitals based solely on the number of inpatient days, however, would not account for the differences in average costs per case arising from differences in the average length of stay (ALOS). The hospital in which the length of stay of patients was longer, other things equal, would tend to have a lower average cost per day because the treatment costs for the additional days would be likely to be far below the average for the case. The extra day's stay in the hospital would probably contribute little to the improvement of the patient's condition, and thus the lower average cost would actually mask inefficient hospital performance.

The figure illustrates the potential use of hospital service statistics in understanding the efficiency implications of unit cost estimates. Knowl-

Figure 3-1. Inpatient Cost Profile



Source: Adapted from Forte, 1985.

edge of the average length of patient stay, the bed occupancy rate (percentage of beds occupied by patients), and the annual bed turnover rate (average number of inpatients per bed during one year) can help in explaining variation in inpatient unit cost measures. Assuming that the treatment cost profile is similar, high occupancy rates tend to result in lower average costs per patient-day because overhead costs are spread over beds that are usually filled. If high occupancy results from relatively few admissions but very long stays, however, hotel costs will be high in relation to the number of patients and average cost per admission will be high. The expected marginal cost per bed-day will be low because the treatment costs at the end of a long hospital stay tend to be minimal. Alternatively, if the bed turnover rate is high, average cost per admission is apt to be lower because hotel costs are spread over a larger number of patients, whereas the marginal cost per day will be relatively high. Increasing the bed occupancy rate through a greater number of admis-

sions per bed rather than longer stays will allow more patients to be served and thus improve hospital productivity.

The service units to which costs are compared in this chapter—admission, patient-day, and outpatient visit—are intermediate output or process measures. Ideally, efficiency in the use of health sector inputs should be assessed in relation to health outcomes (for example, quality adjusted years of life gained, as discussed in chapter 2) rather than process measures. Unfortunately, a massive level of resources would be needed to conduct empirical studies of hospital costs per QALY or other outcome measure for all hospital services, and we are unaware of any such studies having been done. Therefore, we are left to focus on these process measures for analysis of hospital efficiency.

The analysis of hospital costs and efficiency in developing countries is a relatively new phenomenon. Although a few accounting studies of average costs were performed in the 1970s and early 1980s (see, for example, the work on Malaysia by Heller [1975] and reviews of earlier studies in Robertson 1985, and Mills 1987), hospital costing has received increased attention in more recent years. Of the studies reviewed in this chapter, only one precedes 1987. The relative newness and dearth of hospital cost studies circumscribes the contribution of this chapter. On the one hand, the chapter pulls previously widely dispersed information together for the first time and insight is thus gained into economic functioning of hospitals in developing countries; on the other hand the data base for the chapter is relatively small compared with what is available in industrial countries, which underlines the need for further cost studies.

Accounting-Based Cost Studies

In this section we review cost estimates derived from a selection of accounting-based studies. Such studies are often termed "unit cost" studies to indicate that they provide estimates of the average cost of a unit of service. Unfortunately, the term "unit cost" has evolved in the language of health planners to refer ambiguously to both average and marginal costs per unit. To a certain extent the ambiguity in the use of the term "unit" has practical roots. By separating costs into relatively fixed components (utilities and some categories of staff, for example) and variable components (examples are drugs, medical supplies, and food), we can approximate marginal costs by average variable cost using accounting methods. Also, if it is reasonable to assume that average cost is invariant for the relevant scale of production, then marginal and average cost will be equal and estimates of change in cost with projected output can be based on the "unit" cost estimates. Thus, average costs can

be used with care to approximate the marginal cost needed to project the recurrent costs of existing and planned hospitals.

There are a number of additional reasons for measuring average cost. Perhaps the foremost is that an examination of the levels and determinants of costs holds out some possibility of providing useful insight into the relative efficiency of hospital operations. Comparisons of average costs between hospitals with similar roles in a country's health system may be useful for assessing individual hospital performance and identifying hospitals whose average costs are far from the norm. Comparisons of average costs of performing the same activity among different levels and categories of hospitals may help to determine hospital development policy and give an approximation of the potential saving (or increased availability of services) to be derived from improving the referral system. A closely related use of average costs in planning is to provide information that can be used to compare the cost-effectiveness of alternative health sector interventions or to provide information needed to calculate the cost of treating particular diseases and the expenditure avoided by prevention. Average costs also provide information that can be used in justifying budgets when the level of government subsidies is set for public hospitals or when reimbursement is made to nongovernmental organizations. Finally, knowledge of recurrent operating costs is needed for the formulation of economically efficient hospital cost-recovery policies.

Methodology of Accounting-Based Studies

We have divided accounting-based average cost studies into two categories. The first develops detailed cost information for individual hospitals using allocational assumptions in what is sometimes designated a "step down procedure" to distribute aggregate costs across departments and functions. The second makes less detailed estimates of hospital average costs based on aggregate central ministry information or aggregate reported hospital statistics and accounting records.

Step down analyses. Until recently, detailed average cost studies were not available for hospitals in developing countries. Such analyses were seen as too difficult because of the lack of general accounting information and not useful because of the probable imprecision of the results. This view was not correct; a number of studies, many employing variations on the step down costing methodology, have recently demonstrated the feasibility and usefulness of doing cost studies in a variety of economic and developmental environments. To date, this costing method has been applied only as an exception. The experience from these applications

suggests, however, that it would be possible to institutionalize this type of accounting methodology at the hospital, regional, or ministerial level.

Step down cost accounting is a disaggregated method of analyzing the costs associated with specific hospital outputs. It is based on scrutiny of the hospital production process to enable the best assignment of costs to the outputs to which they are related. All hospital expenditures are attributed to specific departments (cost centers), and then allocational criteria, such as time use, are employed to distribute all costs (including overhead and the cost of intermediate outputs) to final service categories. A summary of the step down procedure is given in appendix 3-1.

Our focus is on the final cost estimates derived from step down cost analyses, but it is worth noting that the analytical process can be as important as the final estimate. In itself the process of attributing costs to cost centers gives hospital managers considerably more useful information than the line item accounting information with which they typically operate. The cost center method enables managers at the hospital and departmental level to know the level of resources they have available to produce services, and this information can be used with measures of departmental output (in the dietary department, for example, a comparison can be made between the number of meals served and expenditures in this cost center) to develop benchmarks against which performance can be measured through time and across facilities. The use of cost centers, thus, also promotes financial accountability of departmental managers. The process of allocating costs across intermediate and overhead departments to the final patient service departments yields much greater understanding of resource flows within the hospital. Finally, a comparison of fully allocated costs with service statistics produces average cost estimates of important performance indicators, such as cost per patient-day and cost per outpatient visit. Average cost estimates derived from hospital step down cost studies are included in the top half of table 3-1.

Other accounting cost studies. An accounting alternative to step down analysis is the use of aggregate data, either for individual hospitals or for groups of hospitals. The calculation of meaningful average costs from aggregate data requires that cost and service information be related with respect to time and institutional and geographic coverage. Although studies based on aggregate data require much less analysis than do step down studies, the importance of data coverage, accuracy, and completeness remains the same. Making an informed assumption of the resource use of other outputs (such as an outpatient visit) from a single output measure (such as an inpatient-day) simplifies the process of relating cost and service information to produce average cost estimates. For example, such a study for Rwandan hospitals (Shepard 1988) assumed that the

resources consumed in four outpatient visits equaled that of one patient-day. A single measure of hospital output, the day-equivalent, was created as the sum of patient-days and one-fourth of outpatient visits. Unit cost per patient-day was calculated as the total cost divided by day-equivalents, and cost per outpatient visit was one-quarter of this.

Cost studies based on aggregate hospital information can be performed in less time but provide fewer details and insights than do the step down studies that allow functional analyses. Nevertheless, the information provided from more aggregate studies can be useful in making comparisons among similar hospitals, or different levels of hospitals, and can serve as a basis for making budgetary allocations to hospitals. Aggregate studies should be supplemented by step down studies that give more details on the relative average cost of inpatient and outpatient services to inform the choice of a day-equivalent measure. The aggregate studies can then be used for a quick approximation of average costs for a comparatively large number of hospitals. Average costs based on aggregate accounting studies are cited in the lower half of table 3-1.

Caveats on the Interpretation of Average Costs

Information on average costs provides one useful input needed for assessing hospital performance. Average cost data alone, however, are not sufficient for reaching definitive conclusions regarding hospital efficiency within a country, and even greater caution is warranted in interpreting results from cross-country studies. Differences in the completeness of the data used in each study and in the health, institutional, and economic environment underlying the estimates place limitations on comparisons. Under ideal circumstances, a study comparing the cost per unit of output for several hospitals would tell us which one provided services with the greatest efficiency (technical or economic). The following conditions would have to be met, at a minimum, however, for the results to be unequivocal:

- The quality of services provided in each facility would have to be the same (or adjusted for) so that costs per an equivalent unit of output were being compared.
- The clinical composition of the patients (the case mix) at each of the facilities would have to be the same (or adjusted for).
- For economic efficiency, the cost information would have to measure the social opportunity cost of resources used, not merely the amounts reported to have been spent.[1]

Without an understanding of differences in quality and case mix across hospitals, the efficiency implications of variation in average costs

Table 3-1. Hospital Recurrent Average Cost Estimates from Accounting Cost Studies, Selected Countries, Selected Hospitals
(1988 U.S. dollars)

Country and year of data	Inpatient cost: Level of hospital	Inpatient cost: Per patient-day	Inpatient cost: Per admission	Inpatient cost: Per bed	Outpatient cost per visit	Number of hospitals in study	Source
Step down studies							
Belize, 1985[a]	II	60.4	370	15,075		1	Raymond and others 1987
	III	42.3	126	4,714		6	
China, 1986	I	9.1	260	3,223		1	Chen 1988; Chen 1987
	II	4.4	87	1,446		1	
Indonesia, 1987	II	15.5			6.2	2	Djuhari and others 1988
	III	8.1			2.3	30	
Jamaica, 1985–86	I	30.7	327	9,419	12.6	2	Kutzin 1989
	II	20.9	140	6,646	8.7	2	
	III	23.0	176	6,699	9.0	1	
Malawi, 1987–88[b]	III	3.0	27	1,264	0.6	6	Mills 1991
Niger, 1986–87[c]	I	6.6	93	2,090	15.9	1	Wong 1989
Papua New Guinea, 1988	I	28.7	286	8,384	6.2	1	JSI 1990
	II	25.8	276	7,549	2.5	4	
	III	26.7	364	5,822	5.3	8	
St. Lucia, 1986–87[d]	II	43.9	311	11,907	19.4	1	Russell, Gwynne, and Trisolini 1988
Other accounting studies							
China, 1986[e]	I	8.7	218	2,993	2.2	8	Barnum 1989
	II	5.7	103	1,807	1.4	11	
	III	4.3	57	1,489	1.1	7	

Colombia, 1978	I, II	56.9	421	16,526	14.2	8	PRIDES 1980
Indonesia, 1985[f]	I	18.6	174	5,102	4.6	2	Barnum 1987
	II	14.6	127	3,647	3.7	15	
	III	7.0	41	1,384	1.8	296	
Rwanda, 1984[f]	I	16.1		5,146	4.0	2	Shepard 1988
	II	13.8		4,218	3.5	1	
	III	7.6		1,572	1.9	17	
Turkey, 1987[g]	I	39.8	372	10,649		3	World Bank 1990
Zimbabwe, 1987[h]	I	28.0	219	9,123	10.6	4	Hecht 1992
	II, III	17.9	109	4,366	1.9	90	
United States, 1988[i]	All acute	581.1	4,194			5,579	AHA 1989

Note: Current estimates in local currency were converted to U.S. dollars by exchange rate in that year. The U.S. consumer price index was used to adjust to 1988 terms.

a. For level III hospitals, attribution of inpatient and outpatient costs was based on relative amounts estimated for Belize City Hospital.

b. Outpatient department unit cost relates to new outpatients (unit cost across all outpatients would be lower).

c. Intermediate service (for example, laboratory) costs were attributed to inpatient and outpatient services for comparability with other studies.

d. Radiology, laboratory, operating theater, and physiotherapy costs were attributed to inpatient and outpatient services for comparability with other studies.

e. Cost data identified more than 50 percent of expenses as either inpatient or outpatient. For the remainder, expenditures were attributed on the assumption that the cost of four outpatient visits was equal to that of one patient-day.

f. Day equivalents assume that the cost of four outpatient visits equal that of one patient-day.

g. Total hospital recurrent costs were divided by inpatient statistics; no attempt was made to apportion costs between inpatient and outpatient.

h. Estimates were based on Hecht's assumption that 80 percent of total hospital costs were attributable to inpatient services. The category "II, III" includes provincial, district, and small rural hospitals, plus health centers.

i. Estimates exclude the value of staff time of physicians who are not salaried employees of hospitals.

Source: As noted in table.

cannot be properly interpreted. For example, high average costs may reflect high quality, poor efficiency, or the characteristics of the patients at one institution in relation to another. Low average costs may be a result of an inadequate provision of drugs, and thus would represent poor quality, not greater efficiency. If information on the quality of services and the case mix of patients is added to cost data, the efficiency implications of average cost information become clearer.

Conclusions from Accounting Studies

With these caveats in mind, it is still possible to glean useful information from table 3-1. Perhaps the most striking aspect of the table is the magnitude of cross-country variation in the average cost estimates. The level of average costs is associated with per capita GNP, because the richer countries in the table (Belize, St. Lucia, Turkey, and Jamaica) have the highest unit costs, and Malawi and Niger, the poorest nations, have the lowest unit costs. Given that personnel usually comprises the largest component of hospital costs (see discussion in chapter 2) and that wage rates are associated with per capita income, this finding is not surprising. It suggests that the appropriate technological mix of inputs in a poor country's hospitals is probably different, that is, would use relatively more personnel, from that which is appropriate in a richer country's hospitals. The extent of the difference depends on the feasibility of substituting labor for nonlabor inputs.

Cross-country variation in recurrent average costs may also reflect differences in the quantity of recurrent inputs used, which may suggest qualitative differences in hospital outputs across countries. These differences between outputs across countries may mean that, for example, a hospital day in St. Lucia bears little resemblance to a day in a Chinese hospital of the same level. Therefore, comparison of the cost of producing a patient-day in each country is difficult to interpret because the outputs may be qualitatively different. Detailed information on the prices and quantities of inputs is needed to interpret more meaningfully the relative cost-effectiveness of hospital services across countries.

A study by Lewis, Sulvetta, and LaForgia (1990) of one tertiary hospital in the Dominican Republic sheds some light on the relation between quality and average costs. The authors collected information on specific clinical practices in the hospital and compared them with norms of clinical treatment established by Dominican physicians that described the inputs needed to provide adequate quality of care for selected diagnoses. Price information gathered during the study was used to estimate the cost of achieving these norms. To reach the norms, current expenditures for the appropriate diagnostic tests and drugs would each have to increase tenfold. This suggests that current estimates of average

costs at this hospital are measures of the resources used to provide care of low quality. Although this finding is only suggestive because it assesses quality as a process rather than as an outcome, it has important implications for the interpretation of average cost estimates. The development and costing of country-specific norms and their comparison with estimated average costs would go a long way toward enabling a hospital's average costs to be adjusted for quality so that the cost per equivalent unit of output could be compared across facilities.

Data in table 3-1 show that, within a country, tertiary hospitals tend to have the highest average costs and that the less technically complex district-level hospitals have the lowest. There are several possible explanations for this. First, the teaching role of most tertiary hospitals contributes to higher costs. Second, tertiary hospitals are intended for treatment of patients with the most complex and severe conditions. Therefore, higher costs per patient might indicate different case mixes across facilities. Such differences reflect appropriate use of tertiary facilities and are not, in themselves, cause for concern. Third, higher average costs in tertiary hospitals might also result from their having more equipment and other resources available than lower-level facilities have. This may lead to the use by tertiary hospitals of a more expensive mix of inputs to treat cases of similar complexity and severity to those in provincial or district hospitals. Such use of resources might be inappropriate and require corrective action with respect to referral policy. Treatment norms and analyses of hospital case mix are needed to inform such policy decisions. Each of these factors—teaching function, case mix complexity, and costlier input mix—contribute to the higher average costs found in tertiary hospitals, with the relative contribution of each varying by hospital.

A common finding of the step down studies was that official government budgetary information and even the financial reports of individual hospitals understate expenditures (and, to a greater degree, costs) by and on behalf of public hospitals. Although this finding has implications for future cost studies (particularly for more aggregate-level analyses), perhaps a more important finding is that those within the hospital who are responsible for allocating resources do not have a complete picture of their available revenues. For example, if the hospital's budget does not include services provided by a regional maintenance unit, hospital administrators will find it difficult to manage maintenance services. These detailed studies point to the need for greater transparency in public hospital accounting systems, so that managers have an accurate depiction of the level of resources with which they are working.

The data in table 3-1 allow an indirect comparison of the step down studies with those that use aggregate data. Although the studies were performed on a different set of hospitals, the estimates of cost per day

for Chinese hospitals and for Indonesian hospitals were similar under the two methods, which suggests that reasonably accurate results can be achieved through use of aggregate costs and day-equivalents when available resources do not allow for step down studies. In the case of the Indonesian studies, however, there are greater differences with regard to outpatient costs between the step down and accounting results.

The three studies that used the day-equivalent method—China (Barnum 1989), Indonesia (Barnum 1987), and Rwanda (Shepard 1988)—each assumed that, for line item categories for which the direct cost was unknown, the cost of one inpatient day was equal to the cost of four outpatient visits. For China, a separate linear regression of total cost on inpatient bed-days and outpatient visits gave a rough confirmation of the 1:4 ratio. For Indonesia, Djuhari and others (1988) gave a rough confirmation of the same ratio for lower-level hospitals but suggested a ratio closer to 1:3 for upper-level hospitals. With few exceptions, however, the step down studies found the cost of an outpatient visit to be greater than one-fourth the cost of an inpatient day. This finding suggests that it may be worthwhile for countries with large numbers of hospitals to perform a few step down studies on hospitals of each type to generate the appropriate day-equivalent measure.

Table 3-2 presents the average costs from table 3-1 as a percentage of each country's per capita GNP to show hospital costs in relation to the level of income in each country. The percent GNPN measure is a good indicator of the burden of hospital costs on the overall economy of a country. For example, this table indicates that, compared with other countries, the average cost of a hospital day and the annual cost associated with a hospital bed in Rwanda and Zimbabwe are high in relation to the overall level of resources. Alternatively, it is evident that in the United States, where the absolute level of hospital costs is high (table 3-1), costs per discharge in relation to this country's level of resources are lower than in many poorer countries.

Service Statistics, Efficiency, and the Demand for Hospital Care

The above discussion suggests that although average cost studies yield useful results, accurate interpretation of the implications of the calculated level of average costs requires data on service indicators. In this section we examine data from a number of countries on three interrelated hospital service indicators—bed occupancy rate, average length of stay, and the bed turnover rate—and describe a methodology developed by Pabón Lasso (1986) for assessing hospital performance based on the simultaneous analysis of these statistics. Table 3-3 presents data on these indicators (plus an additional indicator showing the relative use of

hospitals for outpatient services) from hospitals in a number of developing and industrial countries.

The bed occupancy rate measures the percentage of total available beds that are occupied by patients.[2] The average length of stay ALOS is the mean number of days from admission to discharge for each inpatient.[3] The bed turnover rate is the average number of inpatient admissions or discharges per bed.[4] Each of these indicators is usually (but not necessarily) defined on an annual basis and can refer to a particular ward, inpatient department, entire hospital, or group of hospitals. Any one of these indicators provides useful information that can help describe the performance of a hospital's inpatient services, but their explanatory power is multiplied when they are used together. Because they are interrelated, knowledge of any two indicators defines the third. In the following subsections we consider each of the service statistics individually and then present a graphical technique using the indicators simultaneously to examine the relative efficiency of hospital performance.

Occupancy and Turnover Rates

The data in table 3-3 indicate that occupancy and turnover rates vary greatly from country to country and between levels of hospitals. For the most part, occupancy rates decrease as the level of the hospital decreases. There are exceptions, such as China and Malawi, where occupancy rates are high at the district level, but in most developing countries low occupancy rates at the district level are an important reflection of economic inefficiency in the hospital sector. Annual bed turnover rates do not show a consistent trend according to the level of hospital. In many countries, turnover is higher in middle- and low-level hospitals than in tertiary hospitals. In these countries, low- and middle-level hospitals are serving a greater number of patients per bed than tertiary facilities. Even if this is the case, however, there may still be room for improving sectoral efficiency by encouraging more patients to use lower-level rather than tertiary hospitals.

Low occupancy rates are a commonly observed problem in many countries, especially in lower-level facilities. Individual facilities have a level of services, usually somewhere in the neighborhood of 85–90 percent occupancy, at which they have been designed to operate most efficiently. In the short run, a relatively small percentage of hospital costs can be varied; most costs are fixed and determined by the scale of the facility and the personnel establishment (the overhead costs shown in figure 3-1). In chapter 2 it was reported that personnel costs, which make up the bulk of fixed cost, represent a range of about 35 to 75 percent of total recurrent costs. The effect of low occupancy is to spread the cost of personnel and other fixed inputs over a smaller number of service units

Table 3-2. Hospital Recurrent Average Cost Estimates from Accounting Cost Studies, Measured as a Percentage of per Capita GNP

Country and year of data	Level of hospital[a]	Inpatient cost: Per patient-day	Inpatient cost: Per admission	Inpatient cost: Per bed	Outpatient cost per visit	Number of hospitals in study	Source
Step down studies							
Belize, 1985[a]	II	4.9	30	1,231		1	Raymond and others 1987
	III	3.5	10	385		6	
China, 1986	I	3.2	90	1,119		1	Chen 1988;
	II	1.5	30	502		1	Chen 1987
Indonesia, 1987	II	3.6			1.4	2	Djuhari and others 1988
	III	1.9			0.5	30	
Jamaica, 1985–86	I	3.7	40	1,148	1.5	2	Kutzin 1989
	II	2.6	17	810	1.1	2	
	III	2.8	21	817	1.1	1	
Malawi, 1987–88[b]	III	1.9	17	806	0.4	6	Mills 1991
Niger, 1986–87[c]	I	2.2	32	710	5.4	1	Wong 1989
Papua New Guinea, 1988	I	3.3	33	962	0.7	1	JSI 1990
	II	3.0	32	866	0.3	4	
	III	3.1	42	668	0.6	8	
St. Lucia, 1986–87[d]	II	3.0	21	808	1.3	1	Russell, Gywnne, and Trisolini 1988
Other accounting studies							
China, 1986[e]	I	3.0	76	1,039	0.8	8	Barnum 1989
	II	2.0	36	627	0.5	11	
	III	1.5	20	517	0.4	7	

Colombia, 1978	I–II	3.4	25	985	0.8	8	PRIDES 1980
Indonesia, 1985[f]	I	2.8	26	756	0.7	2	Barnum 1987
	II	2.2	19	540	0.5	15	
	III	1.0	6	205	0.3	296	
Rwanda, 1984[f]	I	5.2		1,667	1.3	2	Shepard 1988
	II	4.5		1,366	1.1	1	
	III	2.5		509	0.6	17	
Turkey, 1987[g]	I	3.1	28	816		3	World Bank 1990
Zimbabwe, 1987[h]	I	4.3	33	1,393	1.6	4	Hecht 1992
	II, III	2.7	17	667	0.3	90	
United States, 1988[i]	All acute	2.0	15			5,579	AHA 1989

Note: Percent GNPN cost figures derived by dividing current year average cost estimates by the country's per capita GNP for that year, both measured in local currency.

a. For Level III hospitals, attribution of inpatient and outpatient costs was based on relative amounts estimated for Belize City Hospital.

b. Outpatient department unit cost relates to new outpatients (unit cost across all outpatients would be lower).

c. Intermediate service (for example, laboratory) costs were attributed to inpatient and outpatient services for comparability with other studies.

d. Radiology, laboratory, operating theater, and physiotherapy costs were attributed to inpatient and outpatient services for comparability with other studies.

e. Cost data identified more than 50 percent of expenses as either inpatient or outpatient. For the remainder, expenditures were attributed on the assumption that the cost of four outpatient visits was equal to that of one patient-day.

f. Day equivalents assume that the cost of four outpatient visits equal that of one patient-day.

g. Total hospital recurrent costs were divided by inpatient statistics; no attempt was made to apportion costs between inpatient and outpatient.

h. Estimates were based on Hecht's assumption that 80 percent of total hospital costs were attributable to inpatient services. The category "II, III" includes provincial, district, and small rural hospitals, plus health centers.

i. Estimates exclude the value of staff time of physicians who are not salaried employees of hospitals.

Source: As noted in table.

Table 3-3. Hospital Service Statistics, Selected Countries, Selected Hospitals

Country and year of data	*Level of hospital*	*Occupancy rate (%)*	*Bed turn-over rate per year*	*Mean length of stay*	*Outpatient visits per bed-day*	*Number of hospitals in study*	*Source*
Argentina, 1980	Acute public	66	20.0	12.0	—	All	
Belize, 1985	II	68	40.7	6.1	1.3	1	Raymond and others 1987
	III	31	37.3	3.0	—	6	
China, 1986	I	94	13.7	25.1	2.1	8	Barnum 1989
	II	86	17.6	17.9	3.6	11	
	III	95	26.1	13.3	3.3	7	
Colombia, 1980	I	73	37.8	7.2	—	9	Pabón Lasso 1986
	II	61	38.7	6.0	—	20	
	III	55	42.8	5.2	—	44	
Ethiopia, 1983–85	Urban	47	14.7	11.8	1.7	6	Donaldson and Dunlop 1987
	Rural	59	29.7	7.2	2.4	13	
Fiji, 1987	I	83	42.5	7.2	2.7	3	
	III	46	47.9	3.5	4.2	19	
Indonesia, 1985	I	75	29.2	9.4	2.5	2	Barnum 1987
	II	68	28.7	8.7	1.8	15	
	III	54	33.6	5.9	2.2	296	
Jamaica, 1985	I	79	35.2	8.2	1.0	5	GOJMOH 1986
	II	84	43.2	7.1	0.6	4	
	III	61	28.6	7.8	0.6	13	
Jordan, 1986	All MOH	71	66.0	3.9	1.7	—	
Korea, 1986	All	60	17.8	12.3	2.3	546	FKMIS 1987
Lesotho, 1985	I	125	50.7	9.0	0.7	1	
	III	129	54.9	8.6	1.1	7	
	Mission	56	19.2	10.7	1.0	9	

Malawi, 1987–88	III	116	47.4	9.0	1.3	6	Mills 1991
Morocco, 1987	All Public	57	20.2	10.3	—	—	Bennis and others 1990
Niger, 1986–87	I	87	22.5	14.1	0.3	1	Wong 1989
Papua New Guinea, 1988	I	80	29.4	9.9	1.4	1	JSI 1990
	II	80	28.1	10.4	2.2	4	
	III	60	16.9	12.9	1.8	8	
Rwanda, 1984	I	88	—	—	0.9	2	Shepard 1988
	II	83	—	—	0.5	1	
	III	57	—	—	2.0	17	
St. Lucia, 1986–87	II	74	38.8	7.0	0.6	1	Russell, Gwynne, and Trisolini 1988
Tanzania, 1989	Mission	81	33.3	8.9	1.5	43	CMBT 1991
Tunisia, 1989	I	76	27.6	10.1	1.0	9	
Turkey, 1987	I	73	28.7	9.3	—	3	
	All MOH	46	25.3	6.6	—	—	
	All hospitals	50	25.5	7.2	—	—	
Zimbabwe, 1987	MOH I	89	41.7	7.8	0.7	4	Hecht 1992
	MOH II	91	54.5	6.1	1.4	8	
	MOH III	76	40.8	6.8	1.9	31	
OECD mean, 1980–83	All	81	16.4	17.9	—	—	OECD 1987
Finland, 1982	All	85	13.9	22.2	—	—	OECD 1987
France, 1983	All	73	18.9	14.1	—	—	OECD 1987
Ireland, 1982	All	80	32.5	9.0	—	—	OECD 1987
Spain, 1981	All	75	18.6	14.6	—	—	OECD 1987
United Kingdom, 1981	All	81	16.0	18.6	—	—	OECD 1987
United States, 1981	All	79	29.0	9.9	—	—	OECD 1987
United States, 1988	Acute	65	33.2	7.2	1.2	5,579	AHA 1989

— Not available.

Source: World Bank sector reviews and appraisal reports, except as noted in table.

and raise the average cost of services. Even if hospital inputs are being used with technical efficiency, low occupancy implies economic inefficiency.

A high bed occupancy rate does not necessarily indicate better hospital performance. Indeed, bed occupancy rates can be too high, in the sense that the volume of services is above the design level of the facility. The implications of high occupancy for average costs and hospital efficiency are ambiguous without information on the other service indicators. The reason for this is that a high occupancy rate may reflect a relatively efficient situation, as when many patients with modest lengths of stay are served (that is, the hospital has a high bed turnover rate), or an inefficient situation, as when the high proportion of filled beds largely results from long lengths of stay. The latter situation is signaled by a low average cost per day but a relatively high average cost per admission. Consider, for example, the level I hospitals (from the non–step down studies in table 3-1) in China and Indonesia. The Chinese hospitals have a lower cost per day but a higher cost per discharge. From table 3-3 it is clear that the Chinese hospitals have a higher occupancy rate, yet the mean length of stay figures suggest that this high occupancy rate does not reflect an efficient situation. A comparison of the turnover rates shows that, at each level, Indonesian hospitals are serving more patients per bed than Chinese hospitals.

There are other reasons why a high occupancy rate does not by itself imply a relatively efficient hospital. For example, with high occupancy rates, scheduling of individual service activities, maintenance, and management becomes more difficult and more costly. The measured average costs from accounting-based studies do not provide sufficient evidence to support or reject the hypothesis that average costs are regularly lower for hospitals with extraordinarily high occupancy rates. More likely is that the quality of services is compromised as staff attention and laboratory and ancillary services are divided among a greater number of admissions that exceeds the hospital's design capacity. In addition, very high occupancy rates may reflect overcrowding, which can facilitate the spread of hospital-acquired infections.

In general, conditions treated in lower-level facilities require simpler interventions and may require shorter lengths of stay than more complex cases seen in tertiary facilities. In those countries where turnover rates are highest in tertiary hospitals, the likelihood is that a large percentage of the cases treated in these hospitals are basic cases not requiring tertiary care. Although the hypothesis needs to be tested in future research, there is a presumption that turnover at tertiary facilities that is high in relation to turnover at lower-level hospitals may be an indicator of sectoral inefficiency.

Quality and Hospital Use

The determinants of low hospital turnover and occupancy rates are not well understood and need careful empirical research to establish statistically confirmed causes. Field experience and qualitative analysis do, however, suggest some likely hypotheses. Possible causes include demand-side factors affecting use that are not within the control of central health authorities, for example, underlying morbidity patterns, reflected in a high proportion of communicable diseases that require less hospitalization than the mix of diseases in higher-income areas, and education and cultural factors, specifically, the lower inclination of poorer and less-educated people to look to hospitals for care. Supply-side factors that interact with demand, however, are probably at least as important and are more amenable to policy intervention. These include cash prices, in the form of user fees, for services and drug charges; nonmonetary prices of access, for example, the value of time spent in gaining access, which is inversely related to the proximity of the hospital to patients in the catchment area, plus the availability and monetary cost of transport; and the quality of services with respect to the adequacy of drugs and other medical supplies, staffing, and the availability of critical specialties.

Quality has both supply- and demand-side characteristics. The critical demand issue is *perceived* quality: the consumer's assessment of the relative quality of different health care providers. Differences in perceived quality, with a basis in fact, provide an important explanation of why some people bypass district facilities and refer themselves directly to provincial or central tertiary facilities despite the greater price in time and money that use of these facilities often entails. Adequate staff and supplies are obvious supply-side factors affecting *actual* quality of services that are important in affecting perceived quality. Thus, demand for services can be responsive to policies on the supply side to improve the availability of key inputs at lower-level facilities.

Availability of drugs and supplies. In low-income countries, a sporadic supply of drugs and supplies, especially in lower-level facilities, is commonplace. Foreign exchange constraints severely limit the purchase of drugs by central ministries, and poor distribution systems further restrict the regular availability of drugs at lower-level facilities in outlying areas. The correlation between drug availability and use of services is quite clear to local hospital managers and is revealed by the variation in occupancy and outpatient visits during the course of a year—when drugs are available service use is high and when drug supplies are limited service use is low. This expected effect of drug availability on perceived quality and thus on demand is supported by a World Bank

study of health facility demand in Nigeria's Ogun State, which found that facility use increased with the percentage of time during the year that drugs were available (Akin and others 1991).

Important to note is that the shortchanging of drug and supply expenditures in hospitals may not result in lower average costs but quite the opposite. The elasticity of service demand with respect to drug availability is probably well above, say, 0.3 at low levels of drug provision, so that a 10 percent fall in drug availability is accompanied by a more than 3 percent fall in service demand. In this case a simple calculation will show that, if drugs make up 30 percent of total hospital unit costs, average costs will increase as the availability of drugs decreases even though drug costs are reduced.[5]

Lack of skilled staff. A similar deficiency in demand is associated with a lack of trained personnel and skilled functions in lower-level hospitals. Because of greater opportunities for private practice, better opportunities for advancement, better amenities, and training that emphasizes Western medical practices and the use of modern technology, medical doctors often resist assignments to rural areas, preferring to be posted in large urban areas. The lack of skilled staff for basic services such as high-risk obstetrics, radiology and laboratory diagnostics, surgery, and pediatrics undermines patient confidence in lower-level facilities. A lack of skilled staff may also indicate an overall insufficiency of other inputs, such as equipment, and an overall inferior quality of services.

Indonesia offers an appropriate context to examine the relation between low occupancy rates and the quality of hospital services. There is wide variation in the occupancy rates in the 296 lower-level hospitals (designated C and D in the Indonesian classification system). Among all class C and D hospitals, half have bed occupancy rates under 50 percent, and one out of six has rates lower than 25 percent. Some insight into the cause of this variation is given by an analysis (reported in the study edited by Prescott [1991]) of the relation between inpatient use and a number of variables, including staffing characteristics. Inpatient use is weakly related to district population size, is higher for urban hospitals, lower outside Java, and is higher when there is a large proportion of the population covered by the compulsory insurance system for government employees (ASKES). But the most important result is the extremely strong effect of staffing, used as a proxy for service quality. Hospitals that offer surgical services had 42 percent higher inpatient use than similar hospitals that do not offer surgery. Adding a specialist doctor (not necessarily a surgeon) would boost use by 83 percent.

These results suggest a strong conclusion: perceived service quality as measured by physician availability is a key determinant of the level of hospital use in Indonesia. Adding a specialist doctor and providing basic

surgical facilities in the lower-level hospitals could increase service quality and quantity dramatically. The total costs in the hospitals would rise, but average costs could fall if use increased. Health system costs may not increase much if physician staff is transferred from urban to rural facilities. Clearly, the problem in implementing this policy is that it is difficult to attract specialist doctors to small outlying hospitals: the financial opportunity cost to specialist doctors at present levels of public sector salaries levels is too great because of the lack of effective demand for supplementary private practice in poor, remote areas. But this analysis strongly suggests that it might be worthwhile to offer a salary supplement to attract specialists to these hospitals. Without a physician and basic facilities the hospital is little more than an empty shell, and the fixed operating costs and sunk investment costs are substantially wasted. Paying market wages for specialists and providing necessary facilities might yield a high payoff. Alternatively, smaller hospitals with very low occupancy rates could be converted into health centers or clinics that provide primary care. Any excess staff or supplies that may result from this conversion could be reallocated to other facilities.

Average Length of Stay

Average length of stay is an important indicator of the efficiency of hospital resource use. There is no reason to conclude that longer stays contribute to higher-quality care; in fact the steady decline during the last thirty years in lengths of stay in most countries belonging to the Organisation for Economic Co-operation and Development (OECD) has occurred at the same time that the technical quality of care in hospitals has improved. Without information about case mix and severity it is difficult to use length of stay as a direct indicator of efficiency for individual hospitals, but stays that are unusually long raise questions regarding efficiency and should provoke a closer search for an explanation of the cause. Differences in the average length of stay among a large number of hospitals of comparable type imply differences in prevailing treatment practices across countries, but again, case mix must be taken into account. Although average length of stay varies across countries, it appears that stays are generally shorter in developing countries than in Europe and are roughly comparable to the average lengths of stay in the United States (see table 3-3). High fertility, however, may bias the average length of stay in developing countries toward short-term confinement, whereas the higher chronic disease rates in the United States and Europe impart the opposite bias. Thus, average lengths of stay adjusted for case mix may actually be higher in many developing countries in relation to industrial countries than is indicated by the aggregate data of table 3-3. The problems arising from cross-country differences in

the definition of a hospital must be noted in comparisons of average lengths of stay. Whether hospitals include long-term as well as acute care beds is particularly relevant to this issue.

China, with extremely long average stays at all hospital levels, is an important example of instances in which a long ALOS reflects technical inefficiency in hospital resource use. In two general hospitals in Shanghai, average stays in 1986–87 for many conditions (for example, hypertension, bronchitis, malignant tumors) were about two to three times as great as in the average OECD hospital in 1980 (Chen 1987 and 1988; OECD 1985). There are many possible reasons for the longer confinements in China, including a lack of alternatives for long-term care, poor scheduling for diagnostics or surgery, and a poor recovery environment in posthospital home care.

The means by which inpatient services are financed can have a direct effect on average length of stay. If patients do not face a monetary price for each day that they stay in the hospital, they have no financial incentive to minimize their length of stay. From the hospital's perspective the financial incentives can be powerful. If a hospital is reimbursed for the costs of an inpatient stay on the basis of a constant per diem fee, they have an incentive to keep patients for lengthy periods. This incentive can be strong; it was shown in figure 3-1 that the last days of a patient stay have low marginal costs, thus making these days the most profitable under this method of hospital reimbursement. Korean and Chinese hospitals are financed in this way. In Korea, the average length of stay was 10.9 days in 1986 (FKMIS 1987) and in China the average length of stay was 19.1 days in 1989 (from a sample of twenty-six hospitals; Barnum 1989). The lower average length of stay is expected in Korean hospitals because the Korean health insurance system includes substantial per diem cost sharing by beneficiaries, whereas most Chinese covered by health insurance have no copayment requirements. A study (Yu 1983) comparing insured with uninsured patients in a Korean hospital from 1978 to 1980 found that the average length of stay was significantly higher for insured patients.

Alternatively, hospital reimbursement policies can serve to limit length of stay. This is the most likely cause of the short average lengths of stay in the United States, where hospitals are increasingly being reimbursed a case-related prospective price for inpatient stays (by private as well as public payers). Under this system, which was implemented in 1984 for the publicly funded Medicare program (serving the acute care needs of the population age sixty-five and older), patients are categorized—according to their diagnosis, procedures received, and other characteristics—into Diagnosis Related Groups (DRGs), and hospitals are reimbursed a fixed price based on the DRG. If the costs of a patient stay are less than this price, the hospital profits. If they are not, the

hospital loses money on that patient. Clearly, under this system hospitals have had an incentive to reduce length of stay (and, some have argued, other resource inputs) per case and to maximize the number of profitable admissions. The average length of stay in nonfederal acute care hospitals for Medicare beneficiaries fell from 10.7 days in 1980 to 8.6 days in 1987 (NCHS 1989). This 20 percent decline bears out the strength of these incentives.

Developing countries have had little experience with case basis pricing, and it is premature to evaluate or recommend policies of this type for low- or middle-income countries. Case-based pricing as it operates in the United States (the DRG system) requires continual updating and monitoring of payment rates and a massive amount of data collection and reporting. These features of the system limit its use in low-income countries at present. Nevertheless, because of the growing need to control costs, less administratively complex versions of case-based pricing may have application. Simplified adaptations of case-based pricing should be evaluated on a pilot basis, particularly in large university or central hospitals. Brazil's social health insurance system has recently begun to reimburse hospitals on a procedure-oriented prospective price basis, and thus one would expect its average length of stay to have fallen since the implementation of this system. Unfortunately, such information is not available (Rodrigues 1989a). In Zaire's Bwamanda Health Zone, the reference hospital is reimbursed a fixed amount per case, with different payment rates for each of sixteen case categories (Shepard, Vian, and Kleinau 1990). As in Brazil, however, no data are yet available to assess the effect of this payment system on ALOS.

Apart from differences in case mix and methods of financing, cross-national variation in average lengths of hospital stay may result from differences in the role of hospitals across countries. In the United States, for example, acute care hospitals rarely provide extended care to those in need of such care. Separate facilities exist to provide long-term care; therefore, the reported mean length of stay of hospitals is not biased by the long stays of those treated in extended care facilities. In many developing countries, separate extended care facilities do not exist, and hospitals must often be a source of long-term as well as acute care (see discussion of extended care, below). In some countries, hospitals also serve other social functions, such as orphanages and as nursing homes for the elderly. These factors must be considered before drawing conclusions regarding the efficiency implications in comparisons of cross-national mean length of stay.

Reducing the average length of stay in upper-level facilities with high occupancy rates would enable turnover rates to increase and thus allow hospital benefits to be extended to a greater number of people, diminishing the pressure for capital investment in new hospital capacity. In

both upper- and lower-level facilities, even where occupancy rates are low, the reduction of excessive ALOS would increase the cost-effectiveness of services by reducing the average cost per admission of specific treatments, although the cost per day may rise from the reduction in relatively inexpensive days. Many factors subject to management intervention have been identified as contributing to long ALOS and are discussed below.

Scheduling. Poor scheduling of diagnostic and therapeutic care contributes to longer stays, especially in upper-level hospitals. In many hospitals it is common for patients to be admitted to inpatient care for diagnostic tests and then confined until results are received. Similarly, patients receiving therapy are often kept in the hospital between treatments rather than treated in short stays or as outpatients. Dramatic examples are the use of scarce high-technology therapies, such as kidney dialysis and radiation treatment, but elective surgery and antibiotic therapy and basic laboratory tests are also poorly scheduled. Blanpain (1987) notes that a selection of Chinese hospitals had waiting periods of three to eleven days from the conclusion of preoperative diagnostics to the day of actual surgery. In the context of figure 3-1 the hospital incurs hotel costs prior to as well as after the principal treatment activity. Insufficient equipment does not fully explain the scheduling bottlenecks; often the same hospitals that report long diagnostic waiting periods have uneven use of special facilities, excess capacity being observed during particular times of the day or week.

Physicians who divide their time between public hospital service and private practice are another source of scheduling problems, although the efficiency of outpatient services may suffer more than that of inpatient services. Many hospitals provide scheduled specialist or nonspecialist services in outpatient clinics. Physicians who maintain a private practice in addition to public employment may experience conflicts in scheduling, which result in their lateness for or absence from scheduled public clinic appointments. It is critical for hospitals to work out arrangements with their physicians to ensure that such conflicts are minimized.

Problems in diagnostic services. These problems often manifest themselves as scheduling problems but result specifically from equipment failures or shortages of supplies or staff needed to conduct diagnostic tests. Lack of recurrent resources to fund maintenance, supplies, and staff adequately is at the root of these problems, which then spill over into delayed testing and, further, into extended lengths of stay. A shortage of reagents often presents itself as a cause of delays in performing and analyzing laboratory tests. Parallel problems arise in radiological services. For example, lack of staff and equipment failure were cited as

the primary reasons behind scheduling delays for radiological exams in Jamaica's public hospitals (GOJMOH 1987).

Extended care. The provision of extended care services in acute care hospitals is sometimes considered an inefficient use of hospital resources. In Europe and the United States, patients convalescing or suffering from long-term degenerative or chronic diseases are often kept in nonhospital facilities (nursing homes, rehabilitation facilities, convalescent homes, or their own homes), where less resource-intensive but more appropriate care is available. Transportation and logistics problems, however, plus shortages of trained staff and other resources, make it difficult to replicate this array of providers in developing countries. Given shortages of recurrent resources for hospitals, many countries probably do not consider the creation of separate facilities for extended care to be feasible. Even though some staff can be reallocated, the additional overhead costs that must be incurred may be prohibitive. Home care, provided reliable outreach support is made available, may be a feasible lower-cost alternative.

The unsatisfactory health environment, poor sanitation, and inadequate housing into which convalescing patients are discharged is often cited by local hospital administrators as a reason to keep patients beyond their direct need for hospital care. Patients or their families may also be a source of demand for longer hospital stays for convalescence or chronic diseases.The extent to which hospitals are used for these purposes probably varies greatly between urban and rural areas and by income level, the poorer and more rural patients using hospitals for extended care to a much lesser extent. Nonhospital alternatives for extended care are discussed further in chapter 6.

Standard treatment practices. Treatment protocols for the same causes of admission vary among countries and through time. After considerable professional controversy concerning the effect of early discharge and the use of ambulatory care on the quality of outcomes, physicians in the industrialized countries concluded that dramatic reductions in length of stay were possible without compromising, and perhaps increasing, the quality of outcome. From 1960 to 1980, the average length of stay was reduced from twenty-one to ten days in the United States (all hospitals; OECD 1987). The reduction was achieved, in part, by changing the standard practice for specific causes of admission. Physicians exert ultimate control over the way in which the treatment plan involves hospitals, including diagnostic and treatment procedures used and lengths of stay. Medical school training and commonly accepted professional practice are determinants of physicians' choices and affect the rate at which new standards are adopted.

Older standards of care and medical protocols may contribute to the longer hospital stays in some countries. To the extent that shorter stays require the use of high-cost modern technology, such change might not be desirable; however, much of the reduced length of stay in OECD countries has been achieved with changed standards based on a better understanding of disease and recovery, changed financial incentives, and widely available modern pharmaceuticals. Through physician retraining programs, government information and management guidance, and appropriate financial incentives (see chapters 4 and 5), practices contributing to shorter stays could be encouraged by government policy choice.

Hospital-acquired complications. Hospital infections and accidents add to extended lengths of stay, but reliable data on rates of hospital infection and their effect on average lengths of stay in developing countries are lacking. Overcrowding, poor sanitation, limited use of disposable items, inadequate use of pressurized steam sterilizers, the misuse of antibiotics, and the use of contaminated blood products all contribute to hospital-acquired medical problems and consequent longer hospital stays. Widespread inadequate maintenance is an important cause of conditions leading to accidents and infection. Policy intervention is also needed to provide explicit guidelines on infection control and to enforce standards of care. A survey of twenty-one hospitals in China measured an overall rate of nosocomial infection of 8.4 percent, with rates as high as 13.2 percent in surgery and 10.3 percent in orthopedics (Blanpain 1987). Despite these seemingly high rates, China's hospital environment may be better than in many other countries.

Matching Facilities to Patient Care Needs

The large average cost differences between high- and low-level hospitals, the inverse relationship between hospital level and occupancy rates, and the scope for further increases of the turnover rate at low-level hospitals indicates that there are economic gains to be obtained from improving the match between patient care needs and the type of facility at which patients seek treatment. An indication of the potential gain from successfully encouraging patients to use low-level rather than middle-level facilities is suggested in table 3-1. Using the averages for the sample of hospitals in Indonesia, and assuming that all cases are of the same complexity and that marginal cost equals average cost, we calculate that the saving (measured in 1988 U.S. dollars) per inpatient bed-day and outpatient visit would be $7.60 and $1.80, respectively. Similarly in Rwanda, the savings per inpatient bed-day and outpatient visit would be $6.20 and $1.60. There are two reasons why these figures provide only

a rough estimate of the potential saving to be gained by such shifts in facility use. First, the average treatment cost of the patients shifted to lower-level facilities is likely to be somewhat less than the average for all patients at mid-level hospitals because the goal is to shift the patients needing basic care toward lower levels, and the average treatment cost of these patients is likely to be less than the average for all patients at mid-level hospitals. The second reason, which perhaps offsets the first, is that marginal costs are likely to be greater than the average cost in mid-level hospitals with high occupancy rates, whereas marginal costs are likely to be lower than average costs in low-level hospitals, which often have lower occupancy rates. Despite these provisos, if patients needing basic care are shifted and appropriate treatment can be provided, they will be treated in facilities with lower overhead costs, and it is likely that a less costly combination of inputs will be used for treatment. This systemwide efficiency gain would provide savings that could be translated into the provision of services to a greater number of clients.

If the patients who have less severe conditions are moved out of tertiary facilities, the remaining case mix in these hospitals will be more costly, on average, than previously, so average costs will probably rise. This is desirable, however, because these hospitals will then be able to concentrate their efforts on the complicated cases for which they were intended. It is expected that average costs in lower-level hospitals will fall, as considerable underused capacity (as evidenced by low occupancy rates) is activated. The magnitude of some of the calculated average cost differences between types of facilities in a country is probably a function of the level of resources available at each level. Successfully encouraging consumers to use lower-level facilities would probably entail increasing the supply of drugs and staff to these hospitals, which might increase their average costs, though they would still likely be less than those in large urban facilities.[6] Nevertheless, a shift of less severely ill patients from crowded urban tertiary hospitals into lesser-used secondary facilities would represent an improvement in resource allocation in the sector.

There is some evidence that self-referral becomes an increasing problem as economic conditions, education, transportation, and communications improve, because these improvements increase the accessibility of higher-level facilities. This phenomenon will occur as long as nothing is done to alter the perception of the quality of services available in basic facilities.

Mechanisms for improving the referral system are related to the quality of services at lower levels and to the financing of services provided at each level. Most important, an effective referral system requires the availability of lower-level services in which consumers have confidence. The means by which services at each level are financed can serve

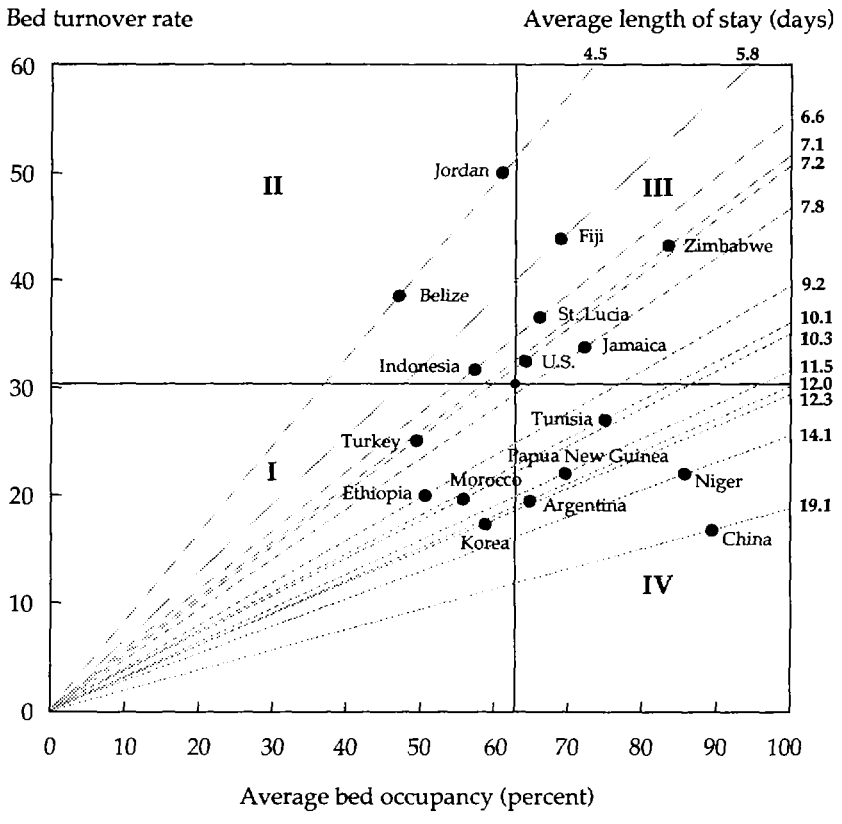
to encourage or discourage appropriate use of the referral system. Achieving a balance between demand and service availability at each hospital level depends on a system of relative prices, fee penalties for nonreferred entry at upper levels, and enforcement of referral that is in balance with the quality of services. Implementation of improved referral policies is facilitated by a district and regional health management structure that includes both primary and secondary care under a common administrative and financial unit responsible for all health facilities and programs in the region.

Simultaneous Use of Hospital Service Indicators to Assess Efficiency

Pabón Lasso (1986) devised a graphical technique to summarize the three inpatient service indicators—occupancy rates, turnover rates, and ALOS—for similar levels of hospitals within a country in order rapidly to assess their relative performance. In figure 3-2 we adapt his graphical technique to describe performance measures across countries. The data represent either the average of a sample of hospitals or the average for all hospitals in the selected countries. The x-axis is the average bed occupancy rate, and the y-axis represents the annual bed turnover rate. Because of the mathematical relationship among these three indicators of hospital performance, a ray drawn from the origin that passes through any point on the graph represents a constant average length of stay, and this measure increases monotonically from left to right across the top and down the right-hand side of the graph.[7]

The graph is divided into four regions by two intersecting lines drawn from the mean values of the bed occupancy and turnover rates (which in turn identify the mean value of the average length of stay). Alternatively, normative values of at least two of these three indicators of hospital performance could be used as the basis for subdividing the graph. Within countries, the division of the graph into four quadrants is useful to identify hospitals that are outliers and demand specific attention. Among countries it is useful to give some idea of the relative performance of the hospital subsector. (The cross-country interpretation needs to be made cautiously because country averages disguise large internal variations.) Pabón Lasso (1986) describes some of the characteristics that could be expected to be found in hospitals falling in each region. As they apply to figure 3-2, hospitals in the countries of region I may be characterized by excess bed availability, low demand for hospitalization in relation to installed capacity, and possibly demand that has been reduced by patients being diverted to other institutions. Hospitals in the countries of region II may have some or all of the following characteristics: excess bed availability, unnecessary hospitalizations, many beds used for patient observation, and a predominance of normal

Figure 3-2. Indicators of Hospital Performance, Selected Countries



Source: See table 3-3.

(as opposed to complicated) deliveries. Region III countries have hospitals that are performing relatively well, on average (though perhaps not compared with normative concepts of what occupancy rates and average lengths of stay should be). They are characterized by a relatively small proportion of unused beds. Hospitals in countries in region IV tend to exhibit some or all of the following: a high proportion of severely ill patients, a predominance of chronic cases, and unnecessarily long inpatient stays.

For international comparisons, this graphical technique is more useful for descriptive than for policy purposes. Figure 3-2 merely describes countrywide averages; it does not suggest specific problems to be ad-

dressed. As an operational tool, it can be applied on a national level as a quick identification guide to hospitals that seem to be performing poorly or very well (see appendix 3-2 for an application to Indonesia). Some of the caveats that apply to the interpretation of unit cost information, however, are also relevant here, especially the need to consider case mix. For example, it would be expected that a tertiary hospital would be clinically composed of more severe cases than would a district hospital and that a greater average length of stay in the tertiary hospital would be one manifestation of this greater severity. The effect of differences in case mix across hospitals can be mitigated somewhat by grouping hospitals according to type and creating one graph per hospital type, as in appendix 3-2. Outliers could then be identified and investigated as to whether they deviate from the norm because of differences in case mix, quality of care, or other "legitimate" reasons, or because of differences in technical or economic efficiency.

Efficiency of Input Use

Inefficiency of input use can lead to high costs per unit of service delivered to patients. In chapter 2 we defined two important sources of inefficiency. The first is technical inefficiency, in which output is less than is technically possible with the mix of inputs used by the hospital (that is, if the institution is operating on the "frontier" of its production possibilities). The second is economic inefficiency, in which the institution may be technically efficient but is not economically efficient in the sense that it is not using the least expensive combination of inputs for its given production of services.

Technical Inefficiency

Technical inefficiency in the use of staff, supplies, and equipment has been cited in many studies of hospital operations. Defining staff assignments too narrowly can restrict the full use of staff time in district-level facilities that have insufficient demand for specialized skills. For example, in a field study of six district hospitals in Malawi, X-ray technicians were used only about half time in their specialty, yet their excess time could not be used in other departments because they lacked qualifications or were unwilling to be reassigned on a part-time basis (Mills, Njoloma, and Chisimbi 1989). This example comes from hospitals with high occupancy, where the effect of technical inefficiency on the quality of health services is likely to be greatest. Nevertheless, the problem of staff inflexibility is more apparent in hospitals with low occupancy rates and high ratios of staff per bed. Broadening the role of lower-level staff to include more support for primary health care activities could add to

the productivity of both hospital and PHC operations. Extending the role of lower-level hospital staff is discussed further in chapter 6.

Theft, bribery, and fraud are serious problems in the management of medical and nonmedical supplies, such as pharmaceuticals and food, in some countries. Security problems can occur at any point in the distribution chain, but hospitals are particularly vulnerable to losses due to petty theft by staff of items for personal use, diversion to the black market, and private medical practices. Poor storage is also a source of pharmaceutical loss in hospitals. Use of careful inventory and accounting systems and well-maintained and secure storage areas is a prerequisite for efficient drug and supplies management.[8]

A large number of sources of technical inefficiency in drug use can be identified. They include extravagant prescribing (a less expensive drug or generic could be used with comparable effect), overprescribing (the drug is not needed or is taken in too large a dose or for too long a period), incorrect prescribing (the wrong drug is prescribed for the diagnosis or it is improperly prepared), multiple prescribing (two or more drugs are used when one would have the same effect), and underprescribing (the dosage or length of treatment is inadequate). The effect of these practices is to increase unnecessarily the cost of effective treatment, as well as to create deleterious health effects. Training, appropriate guidelines, and management surveillance are needed to reduce technical inefficiency in drug use.

Low levels of use and incorrect use of equipment are also sources of technical (and economic) inefficiency at all levels of hospitals. In some cases, especially in low- and middle-level hospitals, the equipment is little used or misused because it is inappropriate. For example, in China, radiological equipment at district and provincial hospitals and computerized axial tomography (CAT) scanners, neonatal intensive care, and dialysis machines at provincial and central hospitals are often unused or inappropriately used (Banta 1987). Fluoroscopes are often used for routine examinations, although they should be reserved for specific problems not resolved by X ray. Expensive 500-milliampere X-ray machines are used rather than the safer and cheaper WHO Basic Radiological System (BRS) (Palmer 1985), which should be all that is needed at the lowest level. In other cases there is a lack of trained personnel to make adequate use of equipment. Ultrasound, for example, is inexpensive and powerful and thus a tempting device to have at low- and middle-level hospitals. Yet reading the ultrasound image requires experience and more training (six months) than often is available.

Another source of technical inefficiency associated with equipment is a lack of uniformity and consistency in supply. A country may receive medical equipment from donations or under multilateral or bilateral assistance programs. Frequently the result is a variety of items from all

over the world and with considerably different technical specifications. Acquiring spare parts and managing the maintenance of, for example, ten different types of X-ray machines from different countries can be extremely difficult. This is an area in which a health ministry must actively manage its donors.

Because all these sources of inefficiency involve a technical waste of resources, it is conceptually possible to obtain more hospital services without any increase in input. Unfortunately, achieving this conceptual possibility can provide a significant challenge; in most cases the increase in output cannot be obtained without deliberate and potentially difficult institutional changes. The inefficiencies that exist have arisen because of the incentives inherent in the existing system of financing health services, a lack of incentives for management or workers, a lack of management skills, a lack of information, or a lack of forethought in investment.

Economic Inefficiency

The production of services, including those of hospitals, involves a choice among alternative techniques and combinations of inputs. Even if the hospital is operating with technical efficiency, hospital managers and service providers are faced with a range of means of achieving a given service level. Ideally, the choice will be economically efficient, that is, the combination chosen to produce a certain quantity and quality of output will have the lowest cost among the range of alternatives. The problem of input choice can be stated as substituting one input for another while keeping the output (quantity and quality of services) the same. Input substitution will occur until the incremental value of services from another unit of expenditure on an input is the same for all inputs. Unfortunately, the lowest cost combination of inputs is often not achieved. The failure to use the economically efficient combination relates to many of the same reasons as the failure to achieve technical efficiency. These include deficiencies in incentives, lack of necessary training or skills, institutional constraints beyond the control of hospital managers (for example, rigid budgetary policies that expressly disallow input substitution), shortages in the flow of real resources to the hospital, and failures in information or investment planning.

Overarching institutional constraints, such as civil service or budgetary regulations, and failures of information and planning must be addressed through coordinated and consistent national policies. Problems of economic inefficiency are generally outside the control of individual hospital managers. It is often at the level of the health ministry that the issues of the appropriate types and quantities of facilities, staff, and equipment must be answered. For example, the least-cost way to improve service coverage may be to convert underused district hospitals

to health centers and reallocate staff. Alternatively, in rural facilities it may be possible to install basic diagnostic equipment that can be operated by lesser-trained technicians and to reallocate more specialized equipment and staff to larger facilities. Analysis of which alternative would achieve the desired population coverage at least cost is not a job for individual hospitals. Regional or national issues of economic efficiency in the health sector involve decisions regarding many issues that affect the performance of individual hospitals but are not within their power to control.

Common problems with economic efficiency can be illustrated by an examination of some substitution choices among important input combinations. In the next few paragraphs we consider the possibilities of substitution between different categories of labor, between equipment or pharmaceuticals and labor, and between maintenance and other expenditure categories.

Substitution among categories of labor. Professional standards of care and accepted practices provide restrictions on service tasks assigned to different categories of medical care staff in a hospital. Most affected are the range of activities undertaken by nurses and paramedical staff compared with the accepted role of medical doctors, as well as the roles within a field or discipline (such as the role of registered nurses compared with that of assistant nurses, or the role of professionally trained radiographers compared with that of X-ray technicians). Possibilities for substitution between and within these categories of staff exist, but they are often limited by roles defined by professional associations.

The quantity of each type of staff person can be related to hospital size and translated into ratios of staff to beds. There are no internationally accepted norms for staffing ratios, and there probably should not be any because staffing choices must be made in the context of local constraints and wage levels. An examination of staffing ratios (table 3-4) reveals cross-country variation in the total staff per bed and, for those countries for which information is available, in the composition of staff. As expected, the limited data support the conclusion that higher-level hospitals tend to employ more staff per bed than lower-level hospitals. In particular, the table supports the observation that physicians are concentrated in large, tertiary hospitals. Nurses and paramedical personnel make up the largest component of hospital staff, although in Indonesia, Jamaica, and Niger the number of other (nonmedical) workers is substantial.

Staffing ratios per bed or bed-day are not an infallible proxy for quality of services. Training and skill level, supporting technology, team work, and the organization of services are all essential complementary co-determinants of quality. In addition, expected differences in case mix

Table 3-4. Hospital Staffing Characteristics and Related Data, Selected Countries, Selected Hospitals

Country and year of data	Level of hospital	Occupancy rate (%)	Bed-days per staff	Staff per bed: Physicians	Staff per bed: Nurses and Parameds	Staff per bed: Other	Staff per bed: Total	Number of hospitals in study	Source
Belize, 1985	II	68	218	0.1	0.9	0.1	1.1	1	Raymond and others 1987
	III	31	225	0.1	0.4	0.0	0.5	6	
Colombia, 1979	I, II	68	100	0.2	1.4	1.0	2.6	8	PRIDES 1980
China, 1986	I	94	177	—	—	—	1.9	8	Barnum 1989
	II	86	172	—	—	—	1.8	11	
	III	95	232	—	—	—	1.5	7	
Dominican Republic, 1989	I	—	—	0.9	0.8	0.4	2.1	1	Lewis, Sulvetta, and LaForgia 1990
Fiji, 1987[a]	I	83	225	0.2	1.0	0.1	1.4	3	
	III	46	176	0.1	0.8	0.1	1.0	19	
Indonesia, 1985	I	75	97	0.6	1.0	1.2	2.8	2	Barnum 1987
	II	68	118	0.4	0.9	0.9	2.1	15	
	III	54	197	0.1	0.6	0.4	1.0	297	
Jamaica, 1985–86[b]	I	84	160	0.2	1.4	0.4	1.9	2	Kutzin 1989
	II	87	238	0.1	0.7	0.5	1.3	2	
	III	80	158	0.1	0.9	0.9	1.8	1	
Niger, 1986–87	I	87	476	0.1	0.3	0.3	0.7	1	Wong 1989
Papua New Guinea, 1988	I	80	328	0.1	0.6	0.2	0.9	1	JSI 1990
	II	80	276	0.1	0.6	0.4	1.1	4	
	III	60	287	0.0	0.4	0.3	0.8	8	

— Not available.

a. Staffing data refer to number of established posts.

b. Staffing data refer to number of established posts as of August 1988.

Source: World bank sector reviews, except as noted in table.

between hospitals suggest that tertiary hospitals require greater staffing intensity than district hospitals. Staffing ratios are, however, an important indicator of hospital performance, and low staffing ratios in district hospitals remain an impediment to adequate service provision in rural areas in many countries.

Substitution of equipment and pharmaceuticals for labor. One reason that ratios of staff to beds vary considerably from country to country, even if quality is roughly comparable, may be the flexibility in the choice of input mix made possible by some forms of new hospital technology. Monitoring equipment can be used in intensive care departments and in wards to replace some nurses, freeing them for other patient care activities. Automated laboratory equipment, such as that for urinalysis and hemoanalysis, can increase the effectiveness and productivity of laboratory workers. The kitchen and laundry can also involve capital-labor substitution possibilities with the introduction of electric, gas, or steam-driven appliances to replace wood fires and hand labor. The potential for efficiency-improving input substitution exists, but it should not be overstated. Although some technological breakthroughs can be characterized as labor-saving, new technology often requires as much labor as older equipment does, and often of a higher level of skill. Country-specific conditions would determine the desirability of such substitution.

Pharmaceuticals can also reduce required staff or increase staff efficiency. Psychiatric drugs render patients more controllable with fewer workers and less-secure facilities. Drugs to reduce infection and increase resistance after surgery can reduce length of stay. The use of drugs, for example, rifampicin or ethambutol in tuberculosis treatment, sometimes makes outpatient care possible for illnesses that previously required long inpatient stays (Barnum 1986).

Substitution between maintenance and other expenditures. Poor maintenance of buildings, equipment, and vehicles is a ubiquitous problem. Maintenance as a share of total recurrent expenditures is often budgeted at less than 4 percent of recurrent costs (for example, see Mills 1991). The level of required maintenance depends on the operating environment and the complexity of the facility, but an estimate of the desirable level of maintenance expenditure for a secondary referral hospital is between 10 and 15 percent of annual recurrent cost.[9] Even if maintenance is budgeted at this level, the actual expenditure is often less, as maintenance is typically the first item to suffer cutbacks in the face of budget shortfalls. Although technically classified as economic inefficiency, underallocation to maintenance has important implications for the overall technical efficiency of the hospital. For example, insufficient mainte-

nance can lead to broken steam units or inoperable X-ray equipment or vehicles, which contribute to the inefficiency of the remaining 96 to 97 percent of recurrent expenditures.

Conclusion. The examples of input substitution given above do not constitute specific recommendations for change. No prejudgment should be made of the appropriate factor mix because what is appropriate will vary across countries and hospitals. For some hospital services, substitution of equipment for labor may reduce diagnostic or therapeutic errors; however, such substitution may also increase errors in other services or in the same services in other hospitals. The choice should be made on the basis of economic efficiency. Careful country-specific analysis needs to be carried out to determine the relative cost-effectiveness of alternative mixes of factors.

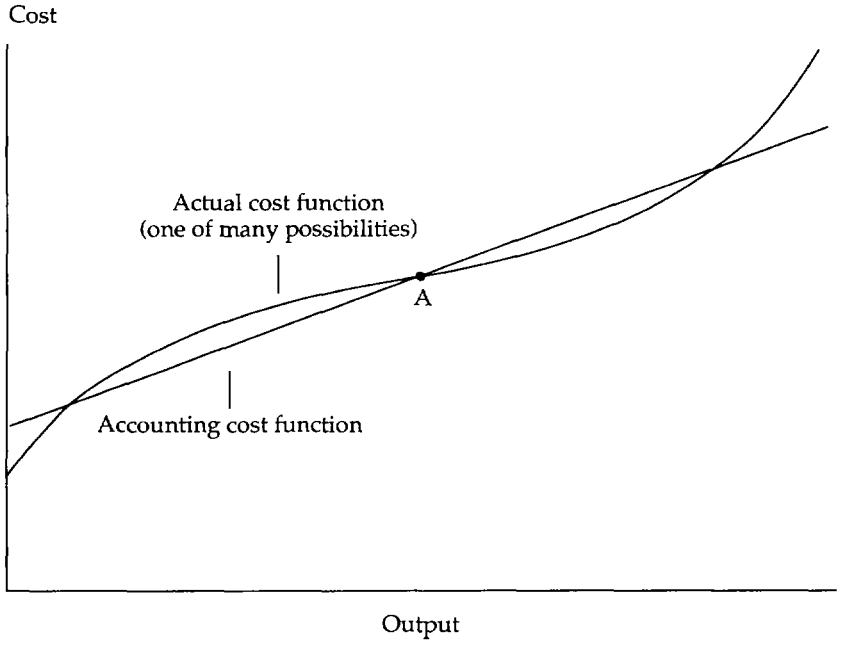
Statistical Cost Studies

The minimum cost of providing a specified set of services when input prices are given can be represented by a point on a cost function. The cost function summarizes the economics of production and can be used to determine the cost of both an additional unit of output (marginal cost) and an average unit and to describe the possibilities of economies or diseconomies of scale or scope. Although hospital managers do not, of course, know the exact nature of their cost function or even of its existence, by their choices through time they necessarily operate within its confines and, if their incentives are sufficiently strong, they operate on or near it. Knowledge of the cost function or at least the rough magnitudes of some of its parameters is especially important for central policymakers because it promotes the setting of policies that are consistent with economic reality. By examining data for a large number of hospitals we can hope to discover some of the characteristics of cost functions. This section summarizes the findings from some statistical cost functions for developing countries.

Modeling Hospital Cost

Accounting studies of costs, such as those reviewed in the first section in this chapter, have an implicit underlying cost function, represented by the sum of the products of the quantity of each input multiplied by its respective price.[10] In fact, such an accounting cost function represents the cost of production, at one point in time, on a cost function. The accounting view of the cost function—sometimes referred to as a "passive" cost function—is rigid and does not allow management or technical responses to changes in input prices or quantities. Figure 3-3 depicts

Figure 3-3. Accounting Cost Function in Contrast to Actual Cost Function



the relation between an accounting cost function and one possibility for the actual underlying cost function. An accounting study generates a point estimate of total costs at the observed output, as at point A in the figure, but does not tell much about what is likely to happen when the price or quantity of an input changes. The only point we are certain that the actual and accounting cost functions have in common, if the hospital is minimizing costs, is the cost and quantity combination depicted at point A. The implicit assumption of an accounting cost function is that the underlying cost curve is linear and, thus, that marginal costs are constant. This assumption may or may not be true and must be tested.

Econometric models of cost and production can be specified sufficiently flexibly that they allow the assumption of constant marginal cost and similar rigid assumptions inherent in the accounting model to be tested. Econometric or statistical models also provide a better depiction of how total costs change in response to differences in service mix, inputs, input prices, and scale of operations. For example, the accounting model would show a direct and simple relation between an increase in a factor price and an increase in total costs; that is, total costs would rise

by the increase in price multiplied by the quantity of the factor. The econometric model would let total costs rise at a rate less than that of the price increase, to allow for substitution of other factors and a reduced quantity of the input for which the price rose. This cost function is curved, not linear as in the accounting model.

In spite of the potential usefulness of econometric cost and production studies, very few such studies of hospitals in developing countries have been carried out thus far. Data deficiencies have been the primary constraint. Few hospitals regularly collect the information needed for such studies, and inconsistencies among hospitals limit the usefulness of the data that are gathered. To a certain extent, statistical methods are robust enough to allow data problems to be absorbed in error terms or imprecision of estimates, but limits on this absorption of poor data quality are soon reached, and the statistical confidence that can be placed in most of the studies to date is low. Another important limitation in econometric studies is that they are built largely on an economic model that assumes the hospital is minimizing costs at some point along its production frontier. Econometric studies also assume that it is fairly easy for managers to change their factor input mix or their capital stock in response to the economic environment. Furthermore, the means by which hospitals are financed have specific implications for the behavior of hospital managers and the characteristics of hospital costs, but the econometric models formulated so far have not been able fully to incorporate the influence of finance. In reality, hospital managers may be insufficiently motivated to operate on the production or cost frontier and instead may be operating comfortably and placidly inside. Or, even if sufficiently motivated, managers may be severely constrained by policies set outside their influence. Finally, as with accounting-based studies, econometric studies suffer from the difficulty of controlling for the quality of output. Thus, the results of the studies to be discussed below are interpreted with caution.

Estimation of hospital cost functions is made difficult by the lack of a clear model of hospital behavior that can be used to interpret the relation between cost and output (service) data. The difficulty of developing an economic view of the hospital lies in the inadequacy of the conventional profit maximization model of the competitive market to explain the incentives, and thus the objectives and behavior, of hospital managers. Many alternative views of hospitals as economic entities have been proposed, ranging from the appealing simplicity of the model suggested by Newhouse (1970)—which proposes that hospitals achieve least-cost production as they strive to maximize a combination of quantity and quality of output constrained by a budget—to a complex, but institutionally apt, model suggested by Harris. Harris (1977) suggests that many hospitals can be characterized as actually composed of two firms—the

first consisting of the medical staff, which delivers the services, and the second consisting of the administrative staff. The input mix, determined in advance by the administrative staff, defines the capabilities of the hospital to meet the demand identified by the medical staff. This view of a hospital may not be appropriate for many hospitals in developing countries, in which the administrative staff and service staff often overlap considerably. It is useful, however, because it recognizes that the demand for services is controlled by the service staff, especially physicians, with resource constraints imposed by managers (centrally or locally). A not uncommon institutional arrangement is that medical staff income is dependent, at least partly, on the volume and type of services offered by the hospital. Thus, this model replaces as a hospital objective the profit maximization of the conventional economic model with the income maximization of physician cum manager.

Still other models have been proposed to explain the incentives of managers in nonprofit organizations in general or hospitals in particular. Pauly (1987) reviews many possible specifications of hospital objectives and notes that, with the current econometric formulations now in use, these models cannot be distinguished from each other. He concludes that the differences in the models may be important for policy decisions involving incentives and hospital pricing behavior, but the models have similar behavioral implications with regard to cost.

A complete review of possible models is potentially large. Indeed, there is no one model that would be appropriate across the tremendous diversity of institutional settings that exist in developing (or industrial) countries. To indicate the variety, a partial list of alternative behavioral models can be given:

- Maximization of output (patient admissions), given a fixed budget
- Maximization of some function of output and quality of care (assuming a tradeoff between the two)
- Minimization of cost, given an exogenous demand for admissions
- Maximization of some function of profit and output
- Maximization of institutional prestige, which is a function of hospital size, facilities, and the prestige of associated physicians
- Nonmaximization—so-called "satisficing" models of behavior—in which managers and staff only hope to achieve some level of output and quality within a fixed budget that will satisfy their own expectations and those of higher-level managers.

Extending Pauly's observations, all these alternatives, with the possible exception of the last two, are consistent with a close empirical relation between total costs and hospital output. The last two models can lead to behavior that obscures the cost-output relationship but, more likely, they

coexist with budget constrained behavior circumscribed by a cost function. Our expectation is, therefore, that even in the absence of an underlying maximization objective that is universally applicable, a functional relation exists between observed hospital costs and output. The relatively close fits of the econometric functions estimated below sustain that expectation.

During the last twenty years, estimation of cost functions for hospitals in industrial countries has become commonplace. Reviews by Cowing, Holtmann, and Powers (1983); Wagstaff (1989a, 1989b), and Breyer (1987) document the progress. The earliest attempts to estimate cost functions using data from hospitals in industrial countries employed specifications of the regression equation that used composite measures of hospital output (for example, Cohen 1967), used average or unit costs of inpatient-day or admission as the dependent variable (Feldstein 1968), and included a variety of interrelated explanatory variables such as occupancy rates, patient flow, length of stay, and capacity as explanatory variables (Mann and Yett 1968). The primary reason the authors of these studies adopted the average cost formulation was to avoid the potential problem of error terms with a nonuniform variance (heteroscedasticity) in the estimated regression. (When total rather than average cost is used as the dependent variable, there is some potential that the error term *may* be correlated with the size of the hospital.) In these estimates, the functional form used in the specification is not derived from a structure based in theory but is generally defined for convenience of estimation as either log-linear or additive-linear.

In more recent attempts, authors have specified a functional form and included variables that are consistent with a theoretical production structure (for example, Cowing and Holtmann 1983; Grannemann, Brown, and Pauly 1986; and Vita 1990). They have generally estimated total cost (rather than average cost) functions with multiple outputs and have employed flexible functional forms. Although the total cost specification can present the problem of heteroscedasticity, this can be dealt with if it arises through the proper selection of an estimation procedure. Some advantages of the more recent work compared with the earlier work are (a) the multiproduct nature of the hospital is explicitly recognized, (b) an economic interpretation of functional form and included variables is provided, (c) econometric problems created by using output (inpatient bed-days) on both sides of the estimating equation are avoided.

Authors of recent cost studies have also been more careful to distinguish between short- and long-run cost functions. This distinction is important for the identification of economies of scale, which is a long-run concept. In empirical applications the distinction between long- and short-run cost functions is related to the specification of the time period

and the inclusion of a scale or capital proxy as an independent variable. If the hospital cannot adjust the capital stock within the period of time defined by the data, as would commonly be the case if the cost and output data refer to a one-year accounting period, then a short-run variable cost function is appropriate. It is possible to examine the behavior of the variable cost function with changes in the capital stock (in this case the number of beds) and derive a long-run function (see Vita 1990). The importance of this is that it is then possible to derive an index of long-run economies of scale from an estimated short-run variable cost function.

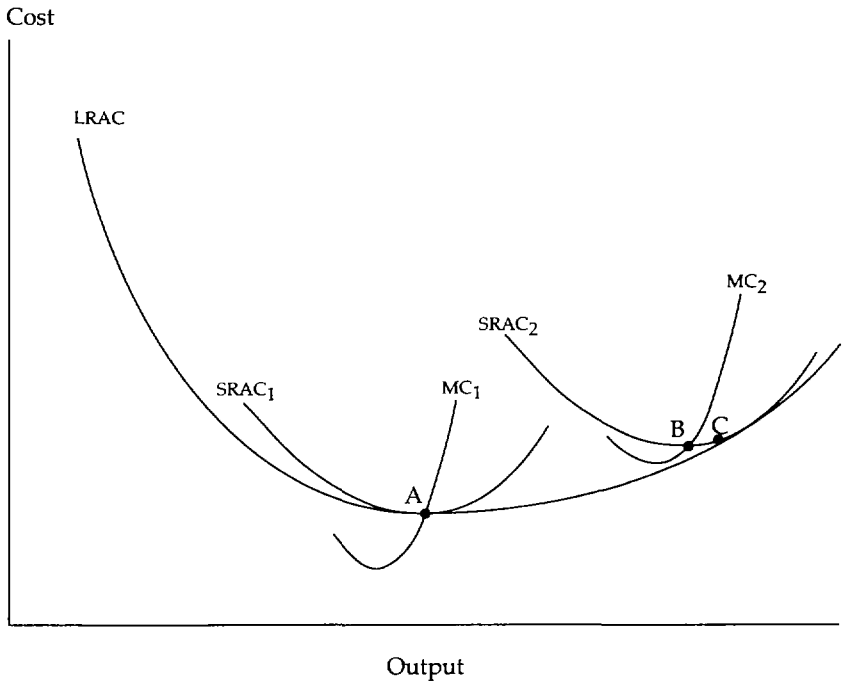
For heuristic purposes (though it may not be descriptive of the actual shape of a given estimated function), we show in figure 3-4 a general graphical description of the relation between long- and short-run cost curves, presented in average cost format, for a single output. With reference to the figure, we can distinguish between the short- and long-run behavior of cost with changes in output by differentiating between returns to scale (a long-run concept) and returns to the variable factor (a short-run concept).

Looking first at different locations along the long-run average cost function (LRAC), we see that the hospital with the short-run average cost (SRAC 1) curve is at an optimum plant size and can be characterized as having constant returns to scale when producing the long-run minimum cost output at A. To the left of A on the LRAC, average cost falls with increases in hospital size (increasing returns to scale). To the right of A, average cost increases with hospital size (decreasing returns to scale).

Looking at locations on a short-run average cost function, for example, SRAC 2, we see that, given the fixed scale, the hospital is at its most efficient level of output at B and is above this short-run optimum at, say, C. At C, the cost of an additional unit—the marginal cost (MC)—is above the average cost. At point B the hospital is at a point of constant returns to the variable factor. To the left of B, short-run average cost decreases with greater output, yielding increasing returns to the variable factor, and to the right of B, say at C, there are decreasing returns to the variable factor. Thus we distinguish between movements along the short-run cost curve by discussing returns to the variable factor and movements along the long-run cost curve by discussing returns to scale. This distinction is not always carefully made in the literature on hospital cost functions. Vita (1990) has recently pointed out that failure to make this distinction has led to errors in estimating economies of scale and that such failure is caused by applying long-run formulas to short-run empirical cost functions. To avoid this problem we explicitly define short- and long-run formulas.

Our central interest in estimated cost functions is in obtaining information on the magnitude of average and marginal cost, on returns to the

Figure 3-4. Relation of Short- and Long-Run Average Cost Functions



variable factor, and on the importance of economies of scale and scope. These concepts can be defined in terms of a general specification of a short-run cost function, $C = f(\text{Beds}, Y_1, Y_2, \ldots Y_n)$, where Beds identifies the scale of the short-run function, C is the total cost of hospital operations, and the Y_i are the hospital's n outputs. We then have the following definitions.

For the short run:

1. The marginal cost of producing an additional unit of the ith output:

$$MC_i = \partial C / \partial Y_i.$$

2. The average incremental cost of the ith output:

$$AIC_i = [C - C(Y_{n-i})]/Y_i$$

where Y_{n-i} is the total cost of production with the exclusion of the ith product. The average incremental cost is the average added cost per unit

of producing the ith product in comparison with producing all products except the ith.

3. Short-run returns to the variable factor:

$$SRVF = C/\Sigma MC_i Y_i.$$

The index of short-run returns to the variable factor measures the effect on costs of a general increase in output when the output mix and bed size remain fixed. If the SRVF is greater than one, the level of output is below optimum efficiency. If the SRVF is less than one the level of output is above maximum efficiency.

4. Short-run product-specific returns to the variable factor:

$$SPRVF_i = AIC_i/MC_i.$$

The indexes of product-specific returns to the variable factor measure the effect on costs of a proportional increase in all inputs on the output of the ith product while the level of output of all other products remains constant. Product-specific returns to the variable factor are said to exist if the SPRVF is greater than one.

5. Economies of scope:

$$SCOPE_S = [C(Y_s) + C(Y_{n-s}) - C(Y)]/C(Y).$$

Economies of scope exist when it is cheaper to produce selected outputs jointly than to do so separately. Economies of scope between a subset of outputs (Y_s) and all other outputs (Y_{n-s}) exist when *SCOPE s* is greater than zero.

In addition to these definitions, we have for the long run:

6. Economies of scale:

With beds included as a proxy for scale, the measure of long-run economies of scale (EOS) is

$$EOS = (1 - \sigma_{C,Beds})/\Sigma\, \sigma_{C,Y_i}$$

where σ_{ab} indicates the elasticity of a with respect to b. The index of economies of scale measures the effect on cost of a general increase in output when the output mix remains unchanged and all inputs are allowed to vary.[11] If the EOS is greater than one, economies of scale are said to exist, if less than one, diseconomies exist.

We now turn to estimated cost functions for developing countries and, where possible, apply these formulas to derive estimates of marginal cost, returns to the variable factor, and economies of scale and scope.

Five Studies of Developing Countries

The literature on developing countries follows much the same evolutionary pattern as that for industrial countries, but in this case there are only a handful of studies. In chronological order the studies known to us cover hospitals in Kenya (Anderson 1980), Ethiopia (Bitran-Dicowsky and Dunlop 1989), Nigeria (Wouters forthcoming). In addition to these we add our own analyses of data for Colombia and China. It is evidence of the relatively recent emergence of this topic that only one of the studies (Anderson's) was obtained from the published literature.[12] Individually, several of these studies suffer from data inadequacies that cloud an interpretation of results; collectively, however, the studies indicate the variation in the magnitude of some critical cost parameters across different countries and different health system environments. We first review the five studies and then briefly discuss the general implications of the findings.

Kenya. The study of hospital costs in Kenya is in the tradition of the earliest studies done in industrial countries and uses an average cost specification:

$$\frac{C}{I} = a_0 \left[\frac{I}{365 \cdot B}\right]^{a_1} \left[\frac{O}{I}\right]^{a_2} \left[\frac{I}{A}\right]^{a_3} B^{a_4} Z_F^{a_5} Z_L^{a_6}$$

where C/I = Costs (C) per inpatient bed–day (I),
$I/(365 \cdot B)$ = Occupancy Rate, B represents beds,
O/I = Outpatient visits (O) per inpatient bed-day,
I/A = Length of stay, A represents admissions,
Z_F = Weighted index of the number of associated subhospital facilities,
Z_L = Level of hospital, 1 = province, 0 = district.

The results of the Kenyan study are clouded by the fact that hospital costs include the costs of related subcenters (health centers and health posts) as well as those of the specified hospital, and it is not possible to match the estimated coefficients clearly with the hospital services per se. An attempt was made to correct for this problem by including a weighted average (Z_F) of the number of subcenters as an explanatory variable. The data set included fifty-one provincial and district public hospitals. Natural logs were taken of all variables before regression so that the estimated coefficients could be interpreted directly as elasticities. The ordinary least-squares estimates of the coefficients are given in table 3-5.

Anderson does not use the estimated equation to calculate the marginal cost of services, and the mean values for the variables employed in the analysis are not reported. He interprets the significant negative coefficient on beds (B) to imply long-run economies of scale with regard

to hospital size, but the specified equation does not yield a direct estimate of economies of scale because of the unit cost formulation and the several places in the specified cost function that inpatient bed-days (*I*) occur. Additionally, Anderson's interpretation of the coefficient on beds is not strictly correct; the estimated function is a short-run cost function, and allowing bed size to change will account for movements from one short-run cost curve to another, keeping output constant, but will not give movements along the long-run cost curve.

By multiplying through by I and rearranging terms, however, we can put the estimated equation in a total cost form. The rearranged equation can be used to derive the elasticities with respect to output and beds that are needed to obtain a measure of long-run economies of scale (EOS above). Following this method, we arrive at 1.5 as the estimate of the index of economies of scale. The policy implications are that, if additional capacity is to be constructed, moderate cost savings could be obtained by expanding existing facilities instead of building new small-scale hospitals. This is, of course, only one of several factors, such as accessibility, that planners would take into account in planning hospital construction.

The function rearranged in total cost form yields an estimate of the index of general short-run returns to the variable factor of 2.0 as well as an estimate of the elasticity of unit cost with respect to inpatient bed-days of –0.8. This would seemingly indicate that Kenyan hospitals can be characterized by decreasing cost with greater output. It would be difficult to reach the conclusion, however, that greater efficiency would be achieved by increasing hospital use and occupancy rates. Anderson points out that many Kenyan hospitals at the time of the study already had occupancy rates in excess of 100 percent and increased use would further compromise quality. Perhaps greater throughput could be achieved by shortening the length of stay; the coefficient on length of

Table 3-5. Cost Function for Selected Hospitals in Kenya

Variable	*Coefficient*	*t-statistic*
Constant	6.57	
I/B	–0.44	–4.17
O/I	0.29	2.45
I/A	–0.07	–0.69
B	–0.20	–2.97
Z_F	0.19	8.23
Z_L	0.29	2.37
	$\overline{R}^2 = 0.75$ n = 51	

Source: Anderson 1980.

stay (I/A) is negative but not statistically significant. It was not possible to calculate economies of scope. A further finding, however, was the significantly greater unit cost for provincial hospitals than for district hospitals. This probably reflects a mix of more severely ill patients in provincial hospitals or a greater allocation of resources per patient to provincial hospitals than to district hospitals.

Ethiopia. As do recent studies of industrial countries, the study of Ethiopian hospitals uses a flexible functional form and estimates a short-run variable total cost function. The specification used is quadratic in outputs:

$$C = e^{(a_0 + a_1 B)}\, e^{f(Y,X)}$$

where

$$f(Y,X) = b_1 I + b_2 I^2 + b_3 O + b_4 O^2 + b_5 I \cdot O + b_6 Del + b_7 Surg + b_8 Lab.$$

All variables remain as defined previously with the addition of the number of deliveries, laboratory tests, and surgical operations represented by Del, Lab, and Surg, respectively. The data set included a pooled cross-section–time-series data set of thirty-eight observations for fifteen public hospitals. Ordinary least-squares estimates of the coefficients are given in table 3-6. The coefficients on outpatients and the square of bed-days are not statistically significant. Taken together, the high $\bar{R}^2$ and large standard errors of the estimated coefficients suggest that multicollinearity is a problem. The sign on the square of bed-days is incorrect and implies an inverted U-shaped average cost curve for bed-days. The peak of the implied average cost curve, however, is well outside the range of the data; throughout the range of the data, average cost increases slowly with increases in output.

The main findings are that, at the sample average, marginal costs are only slightly above average costs for inpatient bed-days and for deliveries and laboratory services. Given the statistical variation in the sample it cannot be clearly established that marginal and average costs differ for inpatient services, and the sample of hospitals is best characterized by nearly constant (slightly decreasing) short-run returns to the variable factor. The study also identifies mild economies of scope between inpatient and outpatient services. Bitran-Dicowsky and Dunlop (1989) do not report estimated results for hospitals of different sizes but note that the output structure varies between large and small hospitals. They suggest that with additional data it would be important to disaggregate the sample by hospital size. Equally important would be a disaggregation by functional level.

Table 3-6. Cost Function for Selected Hospitals in Ethiopia

Variable	*Coefficient*	*t-statistic*
Constant	5.45	22.51
B	4.71 E-3	8.89
I	2.18 E-5	3.44
I^2	−1.65 E-12	−0.02
O	1.91 E-6	0.08
O^2	1.42 E-10	0.26
$I \cdot O$	−7.50 E-10	−2.42
Deliv	1.68 E-4	5.39
Surg	3.21 E-5	0.11
Lab	7.63 E-6	7.97
	$\overline{R}^2 = 0.96$ n = 38	

Source: Bitran-Dicowsky and Dunlop 1989.

Nigeria. The analysis in Nigeria covers twenty-four health institutions in Ogun State, of which eight are health centers with ten to twenty inpatient beds each, seven are maternities (about seven beds each), and nine are dispensaries. Three of these facilities (one health center and two maternities) are privately owned. For our purposes the lack of facilities classified as hospitals limits the applicability of the results. Interpretation of the results is also clouded by the small sample size. In spite of these limitations the study is of interest because it gives some information on the behavior of costs of inpatient care in small facilities and because it uses an econometric specification that explicitly recognizes that facilities may not be operating efficiently. In this technique the production function, designated a frontier production function, is estimated as an envelope of observed points. Wouters (forthcoming) estimates a production function for outpatient visits to compute the marginal products of health workers and non–health workers and compares the ratio of marginal products with the ratio of wages to derive an efficiency index (a value of one represents optimum economic efficiency). The efficiency index is then used in a multiproduct cost function as a technically neutral shift variable. The estimated equation is log-linear. The specified cost equation is,

$$C = a_0\, O^{a_1}\, Z^{a_2}\, W_H^{a_3}\, W_N^{a_4}\, E^{a_5} \exp^{f(A,B)}$$

where

$$f(A,B) = b_0 + b_1 A^* + b_2 B^* + b_3 D$$

and all variables remain as defined previously with the addition of

Z = an index of drug availability as a measure of quality,
W_H = wage of health workers,
W_N = wage of a non–health worker,
E = Efficiency index based on first stage estimate of a production function,
D = dummy variable set to 0 if facility has no beds,
* indicates a Box-Cox transformation ($\lambda = 0.1$) of the variable.

The results of the cost function estimation are given in table 3-7. The main finding from the production function estimate (not reported here) is that the Ogun State facilities use substantially greater numbers of non–health workers than is economically efficient; the ratio of the marginal product of non–health workers to that of health workers is about two-thirds of their ratio of wages. The main findings from the estimated cost function are short-run returns to the variable factor for admissions and approximately constant returns with respect to outpatient visits for the group of small facilities included in the analysis. That is, the marginal cost of admissions is substantially less than the average cost, whereas marginal and average costs for outpatient services are approximately equal. The study also indicates slightly decreasing short-run returns to the variable factor with respect to a neutral expansion of facility production. Long-run economies of scale were not calculated.

Colombia. The Colombian cost function is derived from data given in a special survey (PRIDES 1980) evaluating the financial management of eight class I and II hospitals (*hospitales de referencia*) distributed throughout the country. The costs, given in 1975 constant pesos, and service data cover the period 1975 to 1978 and provide a pooled cross-section and time series data set of thirty-five observations. The function estimated is a short-run variable cost function and includes beds as a measure of

Table 3-7. Cost Function for Selected Health Facilities in Ogun State, Nigeria

Variable	*Coefficient*	*t-statistic*
Constant	1.628	0.58
B^*	0.093	0.44
A^*	0.011	0.25
O	0.597	5.28
Z	–1.361	–2.06
W_H	0.586	2.46
W_N	0.387	2.08
E	–0.160	–0.88
D	0.221	0.62
	$\bar{R}^2 = 0.91$ n = 24	

Source: Wouters forthcoming.

scale. A functional form similar to that used by Grannemann, Brown, and Pauly (1986) was adopted for the analysis. The specified function is cubic in outputs:

$$C = e^{(a_0 + a_1 B)} \cdot e^{f(Y,X)},$$

where

$$f(Y,X) = b_1 I + b_2 I^2 + b_3 I^3 + b_4 O + b_5 O^2 + b_6 O^3 + b_7 I \cdot O$$

and all variables remain as defined above.

A preliminary estimation of this equation gave coefficients on the squared and cubed number of outpatient visits that were not statistically significant and of the wrong sign. Cubic equations often provide difficulty in estimation because of multicollinearity and the magnified effect of outliers on the squared and cubed terms. Reestimation of the cost function omitting the power terms on outpatients gives the results presented in table 3-8. The significance of the coefficient on outpatients remains low, but the sign and magnitude are plausible.

With regard to inpatient costs, the findings are similar to those for the Ethiopian study. At the sample average, marginal costs for inpatient services are approximately equal to average costs, and the sample of hospitals exhibits constant short-run returns to the variable factor. With regard to outpatient services, marginal costs are again very close to average costs and there appear to be constant returns to the variable factor, although in this case the statistical weakness of the coefficient on outpatients reduces the confidence that can be placed in this result. The estimated equation also does not confirm the existence of economies of scope between inpatient and outpatient services. Calculated at the

Table 3-8. Cost Function for Selected Hospitals in Colombia

Variable	*Coefficient*	*t-statistic*
Constant	2.76	10.34
B	3.96 E-4	1.16
I	1.73 E-5	3.27
I^2	–6.42 E-11	–1.74
I^3	8.78 E-17	1.44
O	3.34 E-06	1.28
$I \cdot O$	–7.01 E-12	–0.39
	$\overline{R}^2 = 0.91$ n = 35	

Source: Authors.

means of the data, the estimated value of EOS is close to 1 and suggests constant returns to scale.

China. As a final study, a short-run variable cost function was estimated for the sample of thirty Chinese hospitals summarized in the discussion of accounting unit cost. The sample includes cost and service data for three years, 1984 through 1986. Cost data were converted to constant 1986 yuan. After exclusion for missing observations, the total sample size was seventy-two. Again, a flexible cost function similar to that used by Grannemann, Brown, and Pauly (1986) and cubic in outputs was used for the estimating equation,

$$C = e^{(a_0 + a_1 B)} e^{f(Y,X)},$$

where

$$f(Y,X) = b_1 I + b_2 O + b_3 I^2 + b_4 I^3 + b_5 O^2 + b_6 O^3 + b_7 I \cdot O + b_8 D_{84} + b_9 D_{85}.$$

The variables remain as previously defined with the addition of dummy variables, D_{84} and D_{85}, set to 1 for years 1984 and 1985. After an analysis of the residuals, a maximum likelihood estimate of a frontier production function specification was rejected in favor of an ordinary least-squares estimate (see Aigner, Lovell, and Schmidt 1977). Also, the residuals were not found to be heteroscedastic with respect to bed size (Goldfield-Quandt test). The ordinary least-squares estimates are summarized in table 3-9. The main findings are diseconomies of scale and only mild economies of scope and short-run inefficiencies in the level of operation with respect to bed-days and outpatients.

It is of interest to examine the returns to the variable factor and economies of scale and scope for different levels of hospitals. Table 3-10 gives the marginal cost, short-run returns to the variable factor, and economies of scope and long-run economies of scale for low-, middle-, and high-level Chinese hospitals. The estimates were obtained by substituting the means for the groups in the estimated cost function. The short-run returns to the variable factor of 0.9 for lower-level hospitals and 0.7 for upper-level hospitals indicate that low- and middle-level hospitals are operating at or slightly above efficient volumes of output, whereas larger hospitals are clearly above an efficient volume of output. The index of long-run scale economies indicates that increases in the size of hospitals is not warranted and that, particularly in the upper level, a smaller scale would be more efficient.

These results are consistent with the high levels of occupancy rates in most of the hospitals included in the sample (and in Chinese hospitals in general). Occupancy rates in China are very high, so output takes place

Table 3.9. Cost Function for Selected Hospitals in China

Variable	*Coefficient*	*t-statistic*
Constant	12.93	71.54
B	6.12 E-4	0.72
I	2.03 E-5	4.63
I^2	–6.04 E-11	–1.90
I^3	1.16 E-16	1.56
O	2.01 E-6	1.49
O^2	–9.65 E-13	–0.35
O^3	14.72 E-18	0.83
$I \cdot O$	–11.51 E-11	–2.71
D_{84}	–0.166	–2.23
D_{85}	–0.111	–1.53
	$\overline{R}^2 = 0.89$ n = 72	

Source: Authors.

well to the right of the optimum output on the cost function, in a region where marginal cost is greater than average cost. A mechanical conclusion would be that unit costs could be reduced by using a larger number of institutions and reducing the average occupancy slightly. This interpretation, however, fails to acknowledge the inordinately long stays in Chinese hospitals. The long stays are largely responsible for high occupancy rates, which would be much more modest if length of stay were managed more efficiently. In fact, the productivity of hospital beds in regard to the number of patients treated (reflected in the turnover rate) in Chinese hospitals is lower than in the hospitals of many other coun-

Table 3-10. Marginal Cost and Economies of Scale and Scope by Level of Hospital from Cost Function for a Sample of Chinese Hospitals
(1986 yuan)

	Lower level	*Middle level*	*Upper level*
Marginal cost			
Bed-day	16.2	20.3	40.1
Outpatient visit	1.5	3.2	1.8
Indices of scale and output			
Short-run SRVF	0.9	0.9	0.7
Long-run EOS	0.8	0.7	0.5
Economies of scope			
Bed-days and outpatient visits	0.4	0.3	0.2

Note: Marginal costs are calculated from the estimated cost function in table 3-9 with all variables set at the means of the indicated group.

tries, despite generally much higher occupancy rates in China (see table 3-3). Given the long stays, reducing the cost per day through expanded bed capacity should not be the focus of policy. Reducing length of stay would be likely to increase costs per day slightly (because relatively less expensive days toward the end of the stay would be reduced), yet there would be overall efficiency gains because more patients would be accommodated and the cost per admission would be sharply reduced. The shortened length of stay could also result in a more efficient level of inpatient use.

Discussion of Statistical Cost Functions

The small number of cost function studies for developing countries that are available for the survey above precludes any clear generalization of results. The great variation in results demonstrates that the conclusions must remain country-specific. In addition, it is important to avoid a mechanical interpretation of the statistical results, such as the finding of inefficiency in high Chinese occupancy rates implied by marginal cost that exceeds average cost. The statistical functions are useful only if one has a knowledge of the context.

The scope of the survey could be expanded to include results from industrial countries, but such an expansion should only be done with caution. The greatly restricted budgets of hospitals in developing countries and the limited capacity for training skilled personnel in medical and nursing schools constrain the production technology choices to a much smaller range of activities than in industrial countries, and the underlying production functions are probably very different, especially in low- and middle-level hospitals. Furthermore, hospital managers will respond differently according to the manner in which hospital services are financed. The results from analyses of industrial countries might be of more relevance in the higher-income South American economies and in the emerging countries of the Pacific rim.

Some similarities in the findings between industrial and developing countries do, however, appear indicative. An important debate running throughout the literature of industrial countries concerns the extent of economies of scale. Significantly, the earlier literature, using the single product, unit cost specification, commonly identified significant economies of scale or, strictly speaking, of size. But these findings were criticized, as we noted earlier, because of the poor theoretical specification of the models upon which they were based. The view of economies of scale changed substantially with the introduction of the more flexible estimation forms and the use of total cost, multiproduct specifications. In particular, some of the later studies found constant or mildly diminishing returns to scale, especially for larger hospitals. The general con-

clusion of Cowing, Holtmann, and Powers (1983, p. 276) in their survey of industrial country studies is that "economies of scale may exist for small hospitals but that moderate and large size hospitals can generally be characterized by constant returns to scale."

The results from the five studies of developing countries surveyed here are consistent with the literature concerning industrial countries. With regard to long-run economies of scale, Anderson, using a unit-cost specification in his study of Kenyan hospitals, found economies of scale. However, we used flexible cost functions in our analyses for Colombia and China and found either diseconomies of scale or constant returns to scale at the sample averages. Table 3-11 summarizes the results from the four flexible function cost estimates. With regard to short-run returns to the variable factor, with the exception of the returns to admissions in the small facilities with low occupancy rates included in the Nigerian study, the four studies found either decreasing or constant returns at the sample averages.

In two of the three studies in which the functional form allowed an index to be computed, there was no clear evidence of economies of scope. The Colombian and Chinese data yielded an index value of 0.2 or less between outpatient and inpatient services. The Ethiopian data suggested that some economies of scope may exist between bed-days and outpatient visits. Computation of the scope index for separate levels of hospitals in the Chinese data set (see table 3-10) suggests that the index increases inversely to the level of hospital and that low- and middle-level hospitals with an index value of 0.3 or more may have slight economies of scope in providing both outpatient and inpatient care. Despite the slight variation found in the Ethiopian data and the Chinese middle-level hospitals, the overall results indicate that economies of scope are not an important factor in planning hospital activities.

These results are also consistent with the literature on industrial countries, although in medium to large hospitals some diseconomies of scope have been identified. In their analysis of U.S. hospitals, Grannemann, Brown, and Pauly (1986) found that there were some diseconomies of scope between outpatient visits and inpatient care—hospitals with larger numbers of inpatients also had a higher unit cost of outpatient visits. Their explanation for this phenomenon may be useful in understanding the interaction of inpatient and outpatient costs in large hospitals in developing countries. They state that the difficulty of coordinating a greater range of services may contribute to higher costs. Also, larger hospitals may have outpatient visits of greater complexity that give rise to longer or more costly inpatient stays. They did not suggest, but it may also be, that larger hospitals have more trained personnel and available techniques that are applied to outpatient services independently of case complexity. In other words, the "quality" of

Table 3-11. Marginal Cost and Economies of Scale and Scope from Statistical Cost Functions

	Study			
Item	*Ethiopia*	*Nigeria*	*China*	*Colombia*
Marginal cost				
In 1988 U.S. dollars				
Bed-day	1.4		7.6	22.0
Outpatient visit		2.9	0.5	9.9
Admission		4.2		
Delivery	89.8			
As a % of per capita GNP a				
Bed-day	0.9		1.8	2.7
Outpatient visit		0.3	0.1	1.2
Admission		0.4		
Delivery	55.9			
Indexes of scale and output				
Short-run SRVF		0.8	0.9	1.1
Long-run EOS			0.7	0.9
Product returns				
Bed-day	1.0		0.7	1.0
Outpatient visit		0.8	0.5	0.8
Admission		2.6		
Deliveries	0.9			
Economies of scope				
Bed-days and outpatient visits	0.4		0.2	0.1

a. In year of data.

Source: Bitran-Dicowsky and Dunlop 1989 (Ethiopia); Wouters forthcoming (Nigeria); authors (China and Colombia).

outpatient care in larger hospitals in terms of resources used per patient may be better (although it is not clear that the health outcome for noncomplex cases is superior; that is, qualitatively different inputs are used, which may or may not lead to qualitatively different health outcomes).

The general lack of marked economies of scope for all hospitals and possible diseconomies of scope for larger hospitals removes an economic argument for expanded outpatient departments in large hospitals. Taken at face value, the estimated results imply that limiting, or perhaps eliminating, the outpatient department of large hospitals and shifting the burden of outpatient care to lower levels of facilities would improve the efficiency of the hospital subsector. Much more detailed study is needed before such a far-reaching policy conclusion is recommended, however. Hospital outpatient services are quite diverse, ranging from emergency trauma care to scheduled specialty treatment and consulta-

tion services. It is unlikely that large hospitals could divest themselves of their casualty and emergency departments, nor is it likely that highly specialized outpatient care requiring expensive equipment and personnel could be moved from a hospital setting, as both the capital and recurrent costs of opening separate outpatient facilities are likely to be prohibitive. A much more feasible policy conclusion is that lower-level facilities should be upgraded to provide basic outpatient services to general outpatients requiring nonspecialized care and to provide follow-up ambulatory services to patients who have received specialized inpatient or outpatient care at a tertiary facility. In other words, first-level referral hospitals should be capable of playing a dual role: that of referring patients to more specialized care and that of providing palliative and follow-up care to these patients after they receive specialized services and are referred back down the provider pyramid to their local institutions.

Summary

The growing number of accounting analyses of hospital costs demonstrates both the feasibility of such studies, even in situations in which information on cost and use is ostensibly lacking, and the usefulness of such analyses. Within countries, the studies can be used to develop performance standards, assist with projections of future resource requirements as demand for hospital services increases, and identify hospitals that require special management attention to improve efficiency. Cost studies can be supplemented by examination of service statistics through use of the technique suggested by Pabón Lasso (1986) to provide further identification of hospitals with substandard performance. Once such hospitals are identified, the analysis of specific efficiency issues can quickly move beyond the realm of economics and into management, organization, and personnel planning. Problems with low turnover and occupancy rates, inappropriate length of stay, and inefficient use of personnel and technical inputs may be solvable if direct changes are made in administrative rules or management decisions. But if inefficiency is widespread, it is probable that corrective policies will involve changing the incentives that guide the behavior of management, physician, and client.

The average cost estimates reviewed in the first section of the chapter show the remarkable similarity across countries in the average cost of a bed-day when measured as a percentage of per capita GNP. With few exceptions the cost per bed-day in a middle-level hospital varies from about 1.3 to 3.0 percent GNPN. The actual constant dollar cost varies by considerably more, however, and this variation is undoubtedly accompanied by broad differences in quality. The variations in quality impede

cross-country comparisons in efficiency of service delivery. Two facts do stand out, however. First, the average cost of lower-level hospitals is nearly always substantially less than that in higher-level referral institutions. The waste of resources resulting from the inappropriate use of provincial and central facilities to treat cases that do not require more specialized technical care can be large. The difference between upper- and lower-level facilities in cost of services gives a rough measure of the resource benefits from improved referral policies, although one must keep in mind that some of the cost variation is likely to be attributable to case mix and quality differences. Second, lower-level hospitals are characterized by low occupancy and turnover rates in many countries. These low rates are commonly related to services that are perceived to be and actually are of low quality. Increased use cannot be achieved merely by new written guidelines for a patient referral system; rather, an adequate supply of inputs at lower levels to improve quality is needed to stimulate demand.

The results of statistical studies performed to date are too limited to provide definitive guidance for policy. Most of the studies have been carried out on small or poorly specified data sets that obscure the conclusions. The potential usefulness of such studies is established, however, by the illumination they could shed on the relation between marginal and average cost and between economies of scale and economies of scope. The setting of hospital prices can be done more efficiently with a knowledge of marginal costs, and economies of scale and scope should be included in hospital design and organization. As experience grows with statistical analysis of hospitals in low-resource environments and with larger hospital data bases and improved specifications, the studies will be of increased usefulness for setting prices and planning hospitals. The individual studies surveyed in this chapter demonstrate the variability that can be expected across countries and the need for doing country-specific studies before assuming constant cost. The passive accounting view of the cost function can be taken as a good first approximation of the relation between output and cost and is certainly preferable to basing planning decisions on rough assumptions. The accounting and true cost functions, however, can be expected to diverge for policies that carry output far from prevailing levels.

The findings in this chapter complement the policy discussion in the subsequent three chapters on pricing and hospital service alternatives. Pricing should be intimately related to costs, which constitute half of the information required for determining optimum pricing levels. The next chapter will add demand and income distribution information to cost information to build suggested guidelines for hospital pricing. With regard to hospital alternatives, the discussion of cost and efficiency in this chapter has suggested that improving referral patterns by directing patients to lower-level hospitals and nonhospital alternatives might pro-

vide savings without sacrificing health status. Chapter 6 outlines innovative and practicable lower-level hospital and nonhospital alternatives.

Appendix 3A. Step Down or Cost Center Analysis of Hospital Costs

Anyone who does detailed step down recurrent cost studies faces two significant problems. The first is a problem that all authors of cost studies must address: cost data may not be directly available for individual hospitals. Multiple sources of budgeting (for example, central as opposed to district-level expenditures) and assorted means of making payment for different line items (for example, the salaries of physicians and nurses may be paid by the central health ministry, other salaries may be paid at the district level or directly by the hospital) make the reconstruction of actual expenditures laborious. The second problem, of particular importance for step down studies, is that cost information may be available only on an aggregate basis for the hospital. The need to reconstruct hospital cost data from multiple sources provides some insight into the problems of resource allocation that confront hospital managers. Often, because of institutional constraints and multiple budgetary sources, they do not have a complete picture of their costs or the level of resources that will be available to them over a period of time.

The first requirement of the estimation process is to get as complete a picture of total recurrent hospital costs as is possible. This means supplementing the hospital line item expenditure data with information on resources used that do not appear in the hospital's budget or financial statement. For example, it is common that the financial statements of public hospitals do not include most expenditures on drugs and medical supplies or on maintenance services provided to the hospital. It is necessary to get such information from central medical stores or other central or regional distributional agencies. It is also necessary to estimate nonfinancial costs, such as depreciation and the value of donated goods and services. Next, because most hospital or health ministry budgets are in a line item format (typical line items are salaries, drugs, other supplies, public utilities, and so on), line item costs must be attributed to cost centers, which reflect specific hospital departments. The specific cost centers vary from study to study, but typically three categories of cost centers are used:

- *Overhead.* These cost centers produce only those services that are consumed by other departments (cost centers) of the hospital, not by patients. Examples include Administration, Housekeeping, Maintenance, and Utilities.

- *Intermediate.* These cost centers produce services that are used by other departments but also provide services directly to patients. Examples include Laboratory, X ray, Operating Theater, and Physiotherapy.
- *Final.* These cost centers provide services directly to patients, not to other departments. Examples are Inpatient and Outpatient, with some studies disaggregating these broad categories into specific departments, such as Medicine, Surgery, Obstetrics-Gynecology, and Pediatrics.

Step down costing can be depicted algebraically as follows:

Let C_{iO} = direct costs in Overhead cost center i;
C_{jI} = direct costs in Intermediate cost center j; and
C_{kF} = direct costs in Final cost center k.

The direct costs are the costs attributed to each cost center prior to their allocation to the cost centers associated with hospital outputs. After the direct costs of each cost center (that is, all of the C_{iO}, C_{jI}, and C_{kF}) are identified, the step down method is applied to allocate all costs to final cost centers. The basis for allocating specific proportions of each cost center's costs to other departments should reflect the consumption of the *source* department resources by the *receiving* department (as an example, the distribution of dietary [source] costs among inpatient departments [receiving] would typically be based on the proportion of total patient-days in each inpatient department). First, direct overhead costs are allocated to all other departments. The bases of allocation of costs from each overhead cost center to the other cost centers are proportions that can be represented as:

α_{ij} = the proportion α of Overhead cost center i's costs "used" by Intermediate cost center j;

α_{ik} = the proportion α of Overhead cost center i's costs "used" by Final cost center k;

ρ_{jk} = the proportion ρ of Intermediate cost center j's costs "used" by Final cost center k;

where

$$\sum_i \alpha_i = 1$$

$$\sum_j \rho_j = 1.$$

In the first step the overhead costs are allocated to intermediate and final cost centers using the allocational proportions α_{ij} and α_{ik}, resulting in the first step allocated costs, C'_{jI} and C'_{kF}.[13] Explicitly,

$$C'_{jI} = C_{jI} + \sum_i \alpha_{ij} \cdot C_{iO}$$

$$C'_{kF} = C_{kF} + \sum_i \alpha_{ik} \cdot C_{iO},$$

where

C'_{jI} = fully allocated costs of Intermediate cost center j; and
C'_{kF} = partially allocated costs of Final cost center k.

Then, in the second step, the allocated ("indirect") costs from the intermediate cost centers in the first step are allocated among the final cost centers using the proportions ρ_{jk}. Explicitly, the fully allocated costs are

$$C''_{kF} = C'_{kF} + \sum_j \rho_{jk} \cdot C'_{jI},$$

where

C''_{kF} = fully allocated costs of Final cost center k.

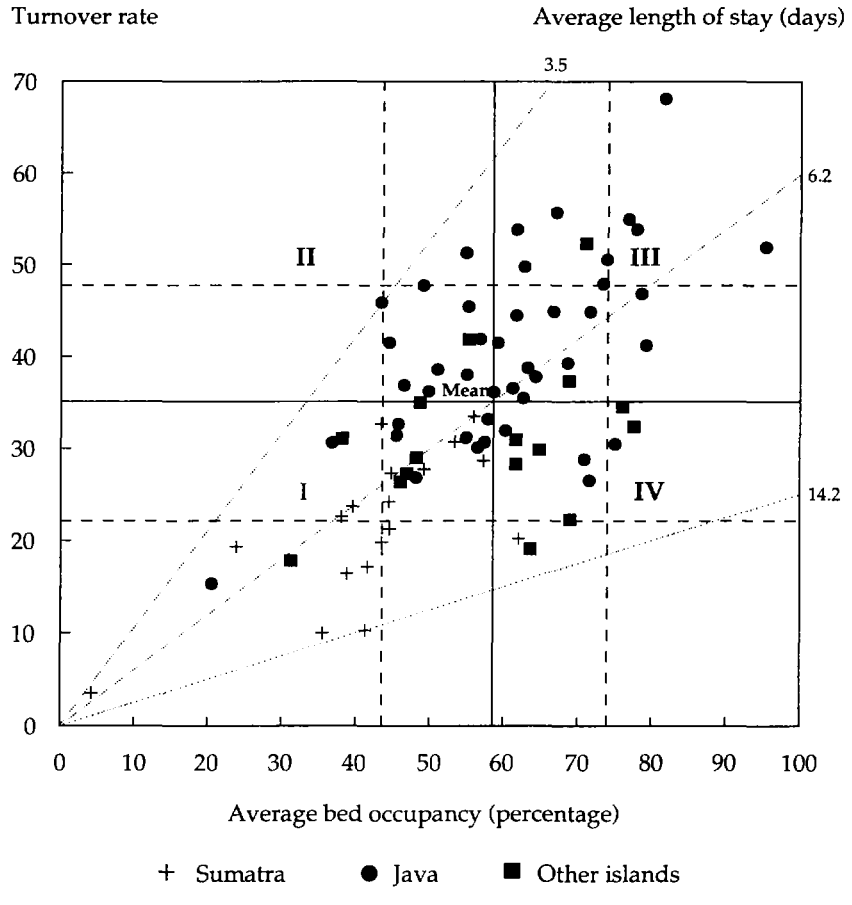
After all costs are fully allocated to each of the final cost centers, average costs are calculated by comparing fully allocated costs to the relevant use statistics. For example, if the only final cost centers are simply Inpatient and Outpatient, statistics on total patient-days and discharges can be compared with fully allocated inpatient costs to generate measures of the average cost per day and per discharge. In a similar manner, the average outpatient cost per visit can be calculated. Average costs of intermediate services can also be calculated if service statistics are available.

Appendix 3B. Use of Service Indicators for Rapid Assessment of Relative Performance of Indonesian Type C Hospitals

Each point in figure 3A-1 (in accord with the method described by Pabón Lasso [1986]) represents one of seventy-eight Indonesian type C hospitals (eighteen on Sumatra, forty-three on Java, and seventeen on the other islands). The points are defined by the rates of bed occupancy and turnover for each hospital; these rates in turn define the average length of stay (shown for its mean and extreme values across the top and down the right side of the graph). The dotted lines are one standard deviation (plus and minus) from the means of the occupancy and turnover rates. Hospitals that lie outside the central rectangle formed by the intersecting dotted lines (with the intersecting solid lines in its center) are considered outliers and merit further investigation to understand their deviation from the norm.

The most striking aspect of the figure is the apparently poor performance of Sumatran hospitals compared with those of Java and the other islands. Most of the outliers in region I are Sumatran. Hospitals in this region have the least desirable characteristics—low use and poor productivity. This region is characterized by low demand for hospital beds in relation to installed capacity, either because of a generalized low demand for inpatient care or because alternatives to type C beds are preferred by the population. The data in the figure suggest that an

Figure 3A-1. Indicators of Hospital Performance in Indonesia, Type C Hospitals, 1985



Source: Barnum 1987.

investigation should be made to explain this situation and develop possible remedies:

- It should be determined whether poor performance is a function of low medical staffing ratios, low budgets per bed, management deficiencies, or overbedding in Sumatra.
- Since demand for type C inpatient services in Sumatra is low, it may be cost-effective to consolidate inpatient services in a smaller number of facilities and to convert some facilities into strictly ambulatory centers. These changes would enable reallocation of staff and allow remaining type C facilities to become more productive.
- If it is determined that demand for inpatient care at a type C hospital is low because the population bypasses this hospital to reach a higher-level (type B) hospital, it may be possible to improve the quality of services visibly at the type C hospital (for example, by transferring resources, such as one or several physicians, from higher levels to the type C) in order to change the population's preferences and improve the referral system. Alternatively, if this is not feasible, it may be better to curtail the inpatient services of the type C hospital.
- The reasons for the relatively good performance of the Javanese hospitals that are outliers in region III should be assessed.

Because of expected variation in case mix, staffing, and possibly other characteristics across hospitals of different types, for example, tertiary in contrast to district, this methodology for assessing hospital performance is most appropriate when applied to facilities that are similar in nature (as in this example). Upon detailed investigation, however, one may find that, even among a group of similar hospitals, deviations from the norm are caused by case mix or other factors that do not necessarily imply differences in relative efficiency across hospitals. Nevertheless, this graphical means of identifying performance outliers among a group of similar hospitals is both quick and effective. Once identified, these outliers can be required to explain their deviations from the norm. By itself, then, this graphical technique does not answer the question of whether a given hospital is performing efficiently. It does, however, enable one to focus attention on those specific hospitals whose performance varies considerably from the norm.

Notes

1. Existing prices may not reflect the true scarcity of certain items in the economy. For example, public sector wage rates may understate the value of physician or nursing services, or an overvalued exchange rate may result in the domestic prices of imported pharmaceuticals being artificially low. Cost estimates should adjust for these distortions, or, at a minimum, note their existence and their approximate magnitude.

2. The occupancy rate (*OCC*) is calculated as:

$$OCC = I / (365 \cdot B), \text{ where}$$

I = Annual number of inpatient days, and
B = Average number of available hospital beds during a year.

3. The average length of stay is defined as:

$$ALOS = I / A, \text{ where}$$

I = Annual number of inpatient days, and
A = Annual number of inpatients (admissions or discharges).

4. The turnover rate (T) is calculated as:

$$T = A / B, \text{ where}$$

A = Annual number of inpatients (admissions or discharges), and
B = Average number of available hospital beds during a year.

5. Let drug availability be measured as drug expenditure (prices are constant) and write total costs as $C = F + D$, where D is drug expenditure, and F is all other costs (assumed fixed in the short run). Define the elasticity of service use with respect to drug availability (expenditure) as $E_{QD} = (\partial Q/\partial D)\cdot(D/Q)$ and the elasticity of costs with respect to drug availability as $E_{CD} = (\partial C/\partial D)\cdot(D/C)$. Average costs can be written $c = C/Q$. Then the change in average cost with a change in drug expenditure is

$$\partial c/\partial D = \left[Q\frac{\partial C}{\partial D} - C\frac{\partial Q}{\partial D}\right]/D^2 .$$

Multiplying by $D/D, C/C, Q/Q$ and rearranging gives the elasticity of average cost with respect to drug availability,

$$\eta_{Cd} = (\partial c/\partial D) \cdot (D/C) = (Q/D^2) \cdot [(\partial C/\partial D) \cdot (D/C) - (\partial Q/\partial D) \cdot (D/Q)]$$

or

$$\eta_{cD} = (Q/D^2) \cdot [E_{CD} - E_{QD}] = (Q/D^2) \cdot [D/C - E_{QD}]$$

where the two terms in the brackets on the right-hand side are the elasticities of total cost and quantity with respect to drugs. If the expression in brackets has a negative absolute value the effect of a reduction in drug availability will be to increase average costs. This highlights D/C and E_{QD} as the appropriate parameters to estimate for an empirical study of the effect of drug availability on demand and average costs. Typically D/C is about 0.3 so that E_{QD} merely has to be greater than 0.3 for a drug shortage to result in increased average costs.

6. The increase would depend on the values of the elasticities in the previous note.

7. On the graph, national mean values of the bed occupancy and turnover rates define a point. Because the average length of stay is defined as

$$ALOS = I/A,$$

knowledge of the occupancy and turnover rates identifies the average length of stay:

$$I = (OCC)(B)(365),$$

and

$$A = (T)(B),$$

therefore

$$ALOS = (OCC)(365) \ / \ T.$$

8. A careful analysis of inefficiencies and remedies for drug supply is given in Quick 1982.

9. In a modestly equipped referral facility, equipment may be 30 percent of total construction cost (about 40 percent of the cost of the building construction). Using a rule of thumb that annual maintenance expenditure should be 1.2 to 2 percent of the capital cost of the building and 7 percent of equipment cost, maintenance should be about 3 percent of total capital cost (0.012 x 0.7 + 0.07 x 0.3 = 0.029). If the recurrent cost ratio (see chapter 2) is between 0.2 and 0.25, maintenance will be between 12 and 15 percent of the recurrent cost (0.03/0.2 = 0.15).

10. The accounting view of the cost function can be written as

$$C = F + \sum_i W_i \cdot L_i + \sum_j q_j \cdot D_j,$$

where F is fixed cost, W and L are the wage and quantity of the ith type of labor, and q and D are the price and quantity of the jth type of other nonlabor inputs.

11. The index of short-run returns to the variable factor, *SRVF*, set out above could be appropriately interpreted as an index of long-run general economies of scale if applied to a long-run cost function in which costs include the cost of capital and the data are defined for a time period that is sufficiently great so that all inputs, including capital stock, can vary. The estimated cost functions presented later in the chapter are all short-run cost functions (as, indeed, are most econometric hospital cost functions). When applied to a short-run variable cost function, the SRVF provides a measure of short-run returns to the variable factor, as we note above, but does not measure economies of scale because scale is fixed. If the function estimated is a short-run variable cost function the appropriate index for economies of scale is the EOS as set out above (Vita 1990).

12. Four of these studies were done to provide background for recent World Bank research. The Ethiopia, Colombia, and China research efforts were done as background studies for this chapter, and the Nigeria study was carried out to support the Nigeria health sector analysis reported in Akin and others (1991).

13. It is also possible to allocate the costs from one overhead cost center (for example, Administration) among other overhead cost centers (for example, Housekeeping). The step down procedure is closely related to input-output analysis of the firm, and it could be formulated in terms of matrix algebra. For an example of step down hospital cost study with a detailed description of methodology, see Russell, Gwynne, and Trisolini 1988.

4. Hospital Financing Alternatives

A growing number of countries have implemented, or are considering, alternatives to government budget allocations for financing health services. Because of the individual nature of the services provided and the importance of hospitals in the total health sector budget, hospitals are a focus for the practical application of these alternatives. Although donors and ministry of finance officials have been supportive and, in some cases promotive, of the use of fees and insurance schemes in the health sector (see, for example, Akin, Birdsall, and de Ferranti 1987), the support has in most cases fallen short of specific guidelines for pricing and revenue collection. In this chapter we summarize the potential role of alternative, nonpublic sources for hospital financing; present a brief outline of the most important issues in financing alternatives; and suggest some general criteria for setting hospital prices.

The impetus for adopting alternative financing policies in hospitals comes from the difficulty of mobilizing sufficient funds for the health sector from public general revenues alone and the inefficiency and inequity of using public funds to support institutions or programs that do not have wide benefits. In addition, problems in achieving or maintaining acceptable quality within government budgets dictate a search for alternative financing sources. Rapid hospital cost escalation has also distorted the allocation of health sector resources between facilities and programs. Financing mechanisms are intrinsically related to potential solutions to these problems because, in addition to augmenting revenues, they affect demand and supply decisions and the allocation of resources.

Broadly, financing alternatives that respond to these problems involve, either separately or in combination, public or private insurance and direct charges to users of hospital services. User fees and insurance are, in themselves, large topics that go beyond the scope of this study.[1] In this chapter and the next, we attempt to keep the focus narrowly on

the financing mode and its effects on the use and provision of public hospital services, although much of the discussion is applicable to other health services as well. In order to clarify criteria for setting health financing policy, we first set out the rationale for public sector involvement in hospitals and identify the objectives of public sector health financing, particularly as these objectives relate to hospitals. Then, in the second section, we discuss alternative financing policies, provide broad principles for designing cost-recovery programs, and briefly outline the institutional characteristics of user fees and insurance as they relate to hospitals.

Health Financing Objectives

Efficiency, equity, and revenue collection are the objectives of a health financing policy. Problems of efficiency and equity in the delivery of hospital services and insufficient funds for recurrent operating costs have been emphasized in the preceding chapters. Hospital financing policy can contribute to improved equity and efficiency of service use and improved funding or, if ill-considered, can contribute to a worsening of these problems. Efficiency, equity, and revenue collection, thus, are criteria by which the performance of a financing policy can be assessed.

Efficiency

The need for public intervention in specific markets, either through provision of goods and services directly by the government or through market regulation, arises from market defects or failures that prevent the achievement of an economically efficient outcome or that lead to inequities that are unacceptable. Not only does the existence of market failures affect the choice of public services to provide, but it can affect the appropriate choice of public financing. In this subsection we briefly review the market failures that provide a rationale for government intervention in the market for hospital services, and in the following subsection we review failures related to equity.

Market failures that lead to allocational inefficiency and that require government financing or provision of services are of great importance for health markets in general and of moderate importance for hospitals. Five general sources of market failure are commonly recognized—public goods, externalities, economies of scale, inadequate consumer information, and incomplete markets. *Public goods* are those goods whose consumption by one person does not diminish the consumption by another and from whose benefits it is not possible to exclude nonpaying consumers if the good is provided at all. *Externalities* are benefits or costs that accrue to persons other than the direct market participants. *Economies of*

scale occur when the unit costs of production decline throughout the range of production. *Information failure* exists when consumers or producers (or both) lack information about the benefits or costs of consumption or production that is needed for rational market participation. *Incomplete markets* occur if uncertainty, the lack of contingent markets (for example, consumer loans), or the lack of future markets prevents consumers and producers from forming a market. Classification of some services as *Merit goods*, is also given as an argument for public sector activities (though this classification cannot strictly be defined as a market failure). The argument for public support of merit goods is not based on economic analysis but is a normative, extramarket, argument based on a subjective evaluation by society (planners or voters) of the desirability of providing and consuming specific goods or services.

Table 4-1 summarizes one view of the strength of each of these sources of market failure as a rationale for government intervention in the provision of selected broad categories of hospital services. The contents of the table are not rigorously determined. The existence or absence of market failure is a question of positive economics and is conceptually

Table 4-1. Strength of Rationale for Public Sector Involvement in Hospital Services

	Rationale					
Type of service	*Public goods*	*Externalities*	*Economies of scale*	*Information failure*	*Incomplete markets*	*Merit goods*
Inpatient						Depends on social and political values
Communicable diseases	..	**	..	**	***	
Chronic diseases	..	*	..	***	***	
Accidents	..	*	*	*	***	
Mental health	..	*	..	*	**	
Surgery	*	*	***	***	**	
Diagnostics	..	*	**	**	**	
Obstetrics						
Normal	..	..	..	*	*	
High risk	..	*	..	**	***	
Outpatient						
Communicable diseases	..	**	..	**	*	
Chronic diseases	..	*	..	***	*	
Emergency and trauma	*	..	**	..	*	
Preventive care	*	***	..	**	..	

*** Strong.
** Moderate.
* Applicable but weak.
.. Of negligible importance.

open to deductive reasoning and empirical validation. Still, there is actually very little related empirical analysis on which to construct such a table, and the importance of the various market failures could also change in different settings. In addition, there is a wide variation of services subsumed within each category. For all these reasons, the contents of the table are open to debate. By citing specific service categories, however, the table provides a concrete frame of reference to allow a judgment of the overall importance of market failure in the case of hospitals.

The first three types of market failures—public goods, externalities, and, possibly, economies of scale—are of importance only in special categories of hospital care. The last two—inadequate consumer information and incomplete markets—are of greater importance. In addition, the merit goods argument, though often implicit, is perhaps the strongest rationale for the public sector provision of hospital services.

With the possible exception of the security provided by the existence of emergency or trauma units, very few hospital services have the characteristics of true public goods. Many hospital services have the weak externality of greater labor productivity of individuals whose health is improved by the services. In addition, a few services, especially those involving communicable diseases, such as tuberculosis, have the important externality of the improved health of others. In general, however, hospital services are characterized by excludability in consumption and the fact that the bulk of benefits accrue to the individual (or household) receiving care. Thus, externalities and public goods are not important in arguments for public provision of services or regulation of the prices of services.

As found in the last chapter, economies of scale are not pronounced in hospitals. When occupancy rates are low, the marginal cost of production may be below average cost, but in hospitals operating at full capacity, marginal cost is likely to be equal to or greater than average cost. It cannot be convincingly argued that hospitals are a naturally declining cost industry that universally require government production, but in smaller communities, especially those in which transportation is limited, an efficient scale of facility may have declining cost in the range of production relevant to the local level of demand. Subsidized production may be required in such communities if the service is to be provided. In addition, some researchers in industrial countries (Grannemann, Brown, and Pauly 1986) have found important economies of scale for emergency care, and the high fixed cost of some diagnostic techniques can also involve economies of scale.

From an efficiency point of view the most obvious reasons for public provision or regulation of hospital care are the inadequacy of consumer information and the incomplete market for hospital services in many

settings, especially in rural and low-income areas. The inadequacy of consumer information restricts the appropriate use of hospital services and creates the opportunity for market exploitation through supplier-induced demand. The consumer has little basis on which to judge the appropriateness of producer-requested surgery, laboratory tests, or other interventions. Public provision of services or private market regulation, although not completely eliminating this problem, can subject producers to greater control. Lack of consumer information may also be an especially large problem in a setting in which the level of education is low, because a failure to recognize the need for care and to seek services at the appropriate time can restrict the demand for services.

On the consumer side of the market, the uncertainty of when diseases will occur, information failure, and the long-term time horizon involved in individual consumption decisions in health lead to incomplete markets for hospital services.[2] Incomplete markets are manifest in the failure of contingency or futures markets to arise to finance private purchases of hospital services. They are especially important in the failure of the market to provide for less frequent and higher-cost medical services even when consumers, if adequate financing were available, would pay to protect against the risk of incurring the full cost of these services. On the supply side of the market, capital market imperfections limit funds for construction of hospitals in rural areas, and skilled labor shortages restrict the operation of rural hospitals.

The market failures cited above prevent the private market from financing an economically efficient allocation of resources directly. Public sector involvement in the market for hospital services does not automatically eliminate the efficiency problems posed by market failures but instead introduces the new problem of selecting an appropriate financing system. The financing system creates incentives for consumers and hospital managers that affect the efficiency of resource allocation. In general, efficiency is served if the price paid by the user reflects the additional benefits to society from consumption of the service and the additional cost to society of producing the service. The appropriate pricing and financing system to achieve this is, however, not obvious. All financing systems have potentially undesirable allocational effects. Regulation or public sector provision of services may lead to nonmarket inefficiency. Hospital services provided without charge may be consumed beyond the point of economic optimality, that is, the marginal cost to society may exceed the marginal benefit of the excess consumption. Regulated prices, as set, for example, by a price control board, may inadvertently provide socially adverse profit incentives for hospital managers to produce particular services, such as CAT scans, and neglect others, such as general wards for inpatients. Finally, unregulated private provision may lead to a service mix that is not socially optimal because

of a lack of consumer information and poorly functioning markets for hospital services. Thus, among the range of choices, from free public provision of the good, through regulated pricing of services provided either by public or private institutions, to a *laissez-faire* private system of service provision, there are associated potential welfare losses. This is a sobering fact, for there are no pat answers to the hospital financing and allocational efficiency problem. The best that health planners can do is to remain aware of the importance of allocational efficiency as a criterion and to attempt to achieve, not necessarily an optimal system, but practical improvements to existing arrangements that are clearly suboptimal, as revealed in the preceding two chapters.

Equity

Evaluating the equity effects of a potential financing mechanism requires identifying who pays, who benefits, and how much (Hoare and Mills 1986). Market failures contributing to a worsening of social equity are of great importance for hospitals. Whereas most publicly provided health care is intended to improve social welfare and involves a redistribution of government revenues collected from a narrow base to provide health services having broad-based benefits, hospitals are, in many countries, an exception. The primarily urban location and the pattern of hospital use by type and cost of service across income classes have equity effects that are too often either neutral or even negative (see discussion of these effects in chapter 2).

As with efficiency, the financing system creates incentives for consumers and hospital managers that affect equity. Financing affects the interaction of supply and demand and thus the distribution of benefits from resources devoted to hospital services. There are also direct implications for equity that arise from government, NGO planning, and entrepreneurial investment decisions. In general, equity is served if the financing system promotes wider accessibility and use of services across income and risk groups. Equity is further improved if the burden of payment is distributed progressively across income. Equity thus involves an interaction of the risks of illness across different social groups, the availability and use of services for the illnesses, and the ability of different groups to pay.

From an equity point of view, incomplete markets for services and the inadequacy of consumer information provide even stronger rationales for the public provision or regulation of hospital services than efficiency does. Information failure is apt to be correlated with low education or rural location, and incomplete markets attributable to limited access to contingency financing are a greater problem for low-income households. By locating in relatively poor geographic areas, using financing methods

that promote the use of services by low-income groups, and providing a referral network that is tied to outreach programs intended to increase accessibility, public hospitals are meant to promote social goals.

In incomplete markets, hospital services are not made available, even if demand would be sufficient to cover costs, because of a structural imperfection that prevents a market from forming. Merit goods are a somewhat different case in that a market with a socially acceptable price and quantity of services cannot be established because of deficient demand (perhaps a result of low household incomes) or the high cost of production. In the absence of government intervention it is unlikely that hospital services would be available and used in many rural areas and among low-income groups in urban areas. Most countries recognize health services as merit goods and provide some services through the public sector, but the scope of the services included varies widely. In some countries, free health services, including hospitals, are specified as a social right. Other countries provide only basic hospital care to indigents as a merit good. Economics does not directly determine what is regarded as a merit good but does determine what is feasible to provide given available resources. Many countries must face the conflict between all that they would like to provide as a merit good and what they can actually afford to provide.

Again, as was true for efficiency, the appropriate financing system to achieve equity goals is not obvious, and all financing systems—whether services are provided free or for fees by the public sector or are provided through the private sector—have potentially deleterious (as well as beneficial) equity effects. The provision of free hospital services by the public sector is potentially the least deleterious, but does have the possible adverse effect of using the public budget to pay for services that do not address the greatest health needs of the poor (either locationally or epidemiologically). Also, if the tax system is regressive and the demand for hospital services is income elastic, the provision of free hospital services may decrease equity. The introduction of fees carries with it the obviously significant risks of damaging the accessibility of health services to the poor. These risks can be reduced if targeting, fee exclusion criteria, and differential pricing can be put into effect. Finally, the equity risk of relying on the private (for profit) provision of services is that such services will cater to the needs of higher-income groups and be offered only where sufficient monetary incentives exist. Conversely, however, each system also has potentially beneficial effects. It is more illuminating to view alternative financing mechanisms not as purely competing but as potential components in an eclectic system of balanced public and private care that combines elements of free service provision, the use of fees, and some form of risk coverage.

Revenue

If we distinguish three levels of budgeting—central finance, the health sector, and the hospital—then we can identify three competing revenue goals that motivate administering agencies to use nongovernmental sources of hospital financing. Conceptually, the three goals result from the desire of the administrative agency at each level to maximize revenues under its control and retain flexibility in allocation across budget items within its jurisdiction. The three competing goals are:

- to supplement a hospital's resources derived from the government budget, that is, to add to the hospital's budget;
- to substitute for the hospital's allocation from the health sector budget and provide supplemental funds for other health activities, that is, to add to the health sector budget and reallocate part of the hospital budget to other programs; and
- to substitute for governmental sources of health revenues, that is, to add to the central government budget.

Outcomes between supplementation and substitution are also possible. For example, a part of alternative collected revenues could be used to supplement the hospital's budgeted resources, a second part could be used to supplement other health activities, and the remainder could provide a partial substitute for the government subsidy to health at either the local or central level. Such a remainder would be analogous to tax revenues and would yield an increase in government general revenues.

These revenue goals and the efficiency and equity goals of cost recovery are related. For analytical clarity, the theory of public finance separates the effects of revenue and expenditure policies, and to the extent practical we attempt to keep this distinction. Yet in execution, revenue and expenditure policies are often linked. The distinction between the three goals above implies that such links are seen by government administrators and motivate revenue policy. If revenues are kept by the hospital and do not supplant existing subsidies, they make improved quality of services or greater quantity of services possible, and the effects on efficiency and equity are then determined by hospital policy. If revenues are kept within the health sector at the local level, they may make additional nonhospital services possible; depending on local health sector policies there may then be cross-subsidization between hospital and nonhospital services and attendant changes in efficiency and equity. The accrual of proceeds of hospital financing alternatives at the central level holds other possible benefits. In this case, the government can use the resources to bolster the central-level program that is deemed to have the greatest social benefit.

Strictly interpreted, the theory of public finance leads to the conclusion that the central level is the best locus for accrued revenues. If geographic redistribution of the funds is needed to improve efficiency and equity, the central government may be a more effective end user of new revenues. Practically, however, accrual of revenues at the local level appears to provide a good compromise between the flexibility of central government accrual and the increase in quality made possible by hospital accrual. In many practical situations, the retention of revenues at the local level would provide the greatest possibility for effective use of the funds. Local governments have the most immediate knowledge of alternative effective and equitable uses of funds and are most flexible in response. Local officials also are more acutely aware of hospital funding requirements. An unconstrained choice among program alternatives, rather than restriction of choice to hospitals or a given sector, can result in the greatest benefit.

Revenue stability. Achieving greater revenue stability for the health sector can also be an objective that motivates administering agencies to use nonbudgetary financing alternatives. During economic recession, brought on, for example, by a decline in the price of an important export commodity such as oil in Indonesia or copper in Zambia, recurrent financing in the health sector falls more or less in tandem with a decline in government revenues. By providing an alternative source of funds that is less sensitive to fluctuations in the government budget, the use of alternative financing for hospitals can reduce the effect of economic recession on health programs.

The efficiency of revenue collection. There is a further relation between these goals, the kind of nongovernmental financing, and the degree of efficiency of collection of nongovernmental revenues (where efficiency is measured as the proportion of potential revenues, given the defined revenue policy, that are actually collected). Depending on the form of financing, the collection of revenues may be more efficient at the central, intermediate governmental, or institutional level. The revenue goals can involve either partial or full recovery of hospital costs (or even the generation of hospital profits). There is a link between the extent to which actual revenue collection can achieve the intended proportion of cost recovered and the kind of financing and institutional form and level of collection.

As an illustration we consider the relation between the efficiency of fee collection and government accrual policy. For any given level of prices, potential levels of demand and revenues are determined. The revenues actually collected from hospital user charges are generally less than those that potentially could be collected if all users of hospital

services paid the designated amount. The principal reason for this is that lax enforcement of fee policy allows nonexempt patients to receive hospital services and never pay for them. Comparing actual with potential revenues allows an assessment of the performance of alternative financing mechanisms, given prices.

Anecdotal evidence (for example, see Collins 1990; Overholt and others 1990; Vogel 1988) indicates that there is an inverse relation between the governmental level at which fees accrue and the efficiency of fee collection. The general recommendation is to allow fees to accrue as closely as possible to the collecting level in order to provide an incentive for managerial surveillance and enforcement of fee policies. There is, of course, no conflict between this recommendation and the revenue goals if the intention is to allow supplementation at the institutional or local level (the first two revenue goals listed earlier). Critically, even if the goal of government fee policy is substitution at the central level (that is, the third revenue goal), and, therefore, a reduction of the net government subsidy after fee revenues returned to the central level are subtracted, some retention of fees at the collecting level is needed to provide a collection incentive.[3]

Alternative Policies

Public sector hospitals in developing countries receive revenues from a large variety of budgeted sources. Financing from government budgets can occur at the central, local, or intermediate level, and at each of these levels more than one agency budget may be involved. Sources of the government budgeted revenues are in themselves based on a diverse tax and financial base, perhaps including earmarked taxes and donor transfers. Collectively, the hospital revenues derived from these sources are a subsidy from the governmental budget allocation. In the next two subsections we discuss fees and risk sharing as alternatives to subsidies from governmental budget allocations that would shift a greater part of the financing burden more directly to household or community sources.

Fees

Fees, if they are to be substantial enough to achieve revenue objectives and cover a significant proportion of hospital costs, must be set with a recognition of their effects on demand and user welfare. Conceptually, fees can improve efficiency of resource use by reducing use of hospital services for care with negligible benefits, removing demand in excess of existing supply capacity, and providing appropriate allocational incentives to both producers and consumers. Fees can also have adverse effects on equity by impeding the access of the poor to needed services.

The few studies that have been done of the demand elasticity of user choice of services (table 4-2) suggest that users are not highly responsive to changes in the price of health care. The low elasticities in these studies, however, do not of themselves indicate either that the efficiency gain from fees or the adverse equity effects of fees is low. First, all the econometric studies on user fees cited in table 4-2 have been carried out where fees are low or nearly negligible. It does not follow that one can extrapolate from the results of these studies to estimate the quantity response to substantial fees. Second, even if the price elasticity is truly low, the welfare effect of fees can be significant because, in households that have fixed incomes, increases in fees imply that the consumption of other goods or services, possibly food or education, could be reduced. Thus, the use of fees can have implications not only for revenue but also for efficiency and equity objectives.

Pricing principles. Guided by the efficiency, equity, and revenue objectives, we can elaborate a normative set of practical principles for a system of health sector fees in government institutions.[4] These principles include considerations of (a) ability to pay, (b) fees as resource allocation signals, (c) the relation between fees and quality, and (d) market failures.

Table 4-2. Price Elasticities of Demand for Health Services

Year of data	*Country*	*Service*	*Price range (U.S. dollars)*	*Price elasticity*
1985	Côte d'Ivoire[a]	Clinic	Free–$0.11	–0.32
			$0.11–$0.22	–0.62
		Hospital	Free–$0.11	–0.38
			$0.11–$0.22	–0.83
1985	Peru[a]	Clinic	Free–$1.56	–0.46
			$1.56–$3.12	–0.68
		Hospital	Free–$1.56	–0.41
			$1.56–$3.12	–0.64
1984	Kenya	Outpatient	Free–$0.13	–0.05–0.20
1975	Malaysia	Public outpatient	—	–0.15
		Public inpatient	—	–0.00
1981	Philippines	Prenatal care	—	–0.01
1985	Ethiopia	Outpatient	—	–0.05–0.50
1990	Nigeria	Outpatient	—	–0.04
1986	Sudan	Outpatient care	—	–0.37

— Not available.

a. Arc elasticity was calculated in price ranges given and for middle-income group.

Sources: Gertler and van der Gaag 1988, 1990 (Côte d'Ivoire and Peru); Heller 1982 (Malaysia); Akin and others 1986 (Philippines); Donaldson and Dunlop 1987 (Ethiopia); Mwabu and Mwangi 1986 (Kenya); Akin and others 1992 (Nigeria); Schwabe, n.d., as quoted in Jimenez 1989 (Sudan).

- *Fees should be consistent with ability to pay and should not prevent essential access to health care.* A direct interpretation of this principle would support price discrimination among users on the basis of income. Some price discrimination may be practical between urban and rural locations or across geographical regions insulated by high travel costs, but within a given geographic location or institution, price discrimination can be difficult to enforce. In the appendix to this chapter, we show that indirect price discrimination in which fees with differing profit or loss margins for services consumed by different income groups can achieve considerable equity gains and can substitute for direct price discrimination.

Detailed price discrimination in which several price tiers are used may be impractical, but some provision should be made to recognize and exempt the very poor from the burden of fees. It may be feasible in smaller institutions to base such exemptions on the judgment of local officials. For larger hospitals, however, a system of formal identification of income status is needed. Still, local control of the identification of the indigent may be practical.

In setting equitable fees, it should be recognized that the actual fee paid to the hospital or clinic is only a part of the true price to the user. Transportation and out-of-pocket drug expenses can be large. In some health systems patients or their relatives supply meals and bed linen. Surveys of the cost of services to patients should inform the design of a fee system. The system of fees needs to provide some limitation on the out-of-pocket expenses of patients. Such a limitation is particularly important for long stays in hospitals or extended outpatient treatment of chronic diseases.

- *Fees should provide correct signals for the direction of the use of health care and health sector resources.* One of the most commonly cited reasons for imposing user fees is to provide signals that discourage unwarranted use of services that have high costs but comparatively low benefits. The system of fees is an essential component in establishing an efficient referral network that avoids the loss of welfare that accompanies inappropriate use of services at upper referral levels by patients who ignore the referral requirements. Patients who use upper-level services directly can be viewed as using luxury services and be charged as such. One possibility is to forgive part of upper-level charges for patients who are appropriately referred.

- *Fees and the quality of services should be linked.* The introduction of fees will result in consumer dissatisfaction and possibly unacceptable decreases in demand if the quality of services is not perceived to be sufficiently high. Quality that is perceived as low at entry levels of referral is the primary reason that clients attempt to avoid established referral chains and go directly to more costly institutions. Several studies have advocated tying increases in fees to increases in quality of services

by using the revenues obtained to improve the availability of pharmaceuticals and provide other improvements in quality. A crucial empirical assumption underlying this recommendation is that the decrease in demand for services in response to the higher fee will be more than offset by the increase in demand in response to the higher quality. A simulation of the equity and efficiency effects of fee increases with accompanying quality increases based on data from several hundred health units in Kenya demonstrated a net increase in welfare (Mwabu and Mwangi 1986). Similarly, simulations based on estimated demand functions demonstrate that quality increases more than offset the minimal reduction in demand that accompanied fee increases in Ogun State in Nigeria (Akin and others 1991). These findings are particularly important with regard to public hospital care at the district level; low usage at this level, even when services are free, is often explained by the poor quality of services and lack of availability of essential complementary inputs.

- *Fees should be subsidized for services that have important externalities, are primarily public goods, have low consumption due to informational deficiencies, or are merit goods.* As noted earlier, goods with externalities have benefits that accrue not only to the individual receiving the service but to others as well. The externalities involved in the consumption of many preventive health services can be substantial and justify government intervention, including subsidization, to achieve an economically efficient level of use. It can be argued that obstetrics and some curative services, such as tuberculosis treatment, have external benefits, but generally hospital services do not have significant externalities. Also, as previously indicated, public goods are characterized by the impossibility of excluding nonpaying users from the benefits of consumption. Examples in health are vector control and some of the community monitoring services of public health laboratories. Hospital services have negligible public good attributes. Inadequate consumer information, in contrast, is an important problem that affects the appropriate use of health care, including hospital services. Without a good understanding of the benefits of a potential service, the consumer cannot determine if or when consumption is warranted. Subsidies are warranted if the lack of information results in underuse of services.

Although merit goods do allow excludability in consumption, it is considered necessary to subsidize them or provide them free to the public because their consumption is deemed to have social merit or be a social right. Public provision of merit goods is related to equity values and political processes. Many countries implicitly consider hospital services merit goods as evidenced by their heavy subsidization, given that they do not qualify for broad-based subsidies on other grounds. Merit goods need continued and careful justification for their public provision, favorable pricing, or subsidization.

Administrative feasibility of fee collection. The introduction of fees in a hospital that has not previously charged for services, or in which the charges have previously been minimal or sporadically collected, can present potentially daunting problems of administration. The management system and physical plant arrangements must provide checks against theft, fraud, and uneven enforcement of exemption rules and allow for accountability and monitoring of the flow of collected funds. Installation and maintenance of a practical system of collection require staff time, training, and even some minimal equipment, all of which have an associated cost. The collected fees (together with the value of any gains in economic or technical efficiency from the introduction of a price system) must exceed the cost of collection in order to justify implementation of a user charge system.

In a survey of fee collection systems in West Africa, Vogel (1988) gives some useful pointers for successful fee collection systems:

- Well-defined entrance points for the hospital
- The issuance of receipts, with duplicate copies, to serve as evidence of payment
- A rigorously enforced system for determining those eligible for exemption
- Training for all staff to confirm the importance of enforcing collection
- Periodic spot checks to establish that the above points are being carried out by all staff
- Periodic audits of the financial transactions and flow of funds.

These elements are needed for successful collection with even the most simply defined fee schedule. Daunting though this may seem, the practicability of fees in diverse settings is demonstrated by the existence of active fee collection in nongovernmental, nonprofit hospitals, often in geographic areas where governmental hospitals provide services without charge.

Optimal pricing. A standard result of economics is the optimality, in the sense of achieving the greatest welfare with a given set of resources, of prices equal to marginal cost. This optimality holds for a competitive equilibrium in markets that supply private goods (as opposed to public or merit goods) to well-informed consumers and in which there is free entry and exit of firms. Hospital services are not provided under these conditions. The failure of the market for hospital services requires regulation of provider behavior and may require subsidized provision of services. Under these conditions the greatest welfare within the constraint of the public budget (or quasi-public institutional budget) can be

achieved by prices that reflect the demand and equity characteristics of the good as well as its marginal costs. The appendix to this chapter reviews some rules of optimal pricing as they relate to the problem of pricing hospital services. The rules provide an explicit specification of prices that are consistent with the general thrust of the principles outlined above. Broadly, the rules contribute to equity goals by incorporating the distribution of income and setting lower prices for services consumed disproportionately by the poor. The rules also contribute to efficiency goals by setting prices that interfere minimally with private preferences.

Detailed exploration of optimal pricing would carry the discussion too far afield and into the realm of technical economics, but it is noted here because it holds out some promise of being useful as a guide in setting rational hospital prices that are consistent with planning objectives. Optimal pricing is suggested in this context, not with the thought that the principle should be applied rigidly, but that it can be used for guidance, together with a less quantitative interpretation of the pricing principles, to set prices that can achieve revenue and efficiency objectives without sacrificing equity.

Price simulations for three broad categories of bundled services, which can be interpreted as successively higher amenity levels of inpatient care, are discussed as an example in the appendix to this chapter. The simulations apply optimal pricing rules, using plausible ranges of required information (on income distribution and distributional objectives, price and income elasticities, budget and subsidy levels) to derive pricing coefficients that are multiplied by the marginal costs of services to give the optimal prices. Some broad implications can be derived from the simulations.

First, the optimal prices are small, but positive, for services that would be used by groups that have the lowest income. Given the plausible sizes of income elasticities used in the simulations, the proportion of marginal costs recovered is very small for the category with the lowest income elasticity (this would correspond, say, to ward care) but substantially higher for the category with the highest income elasticity (this would correspond, say, to a private room with special amenities). In hospitals with 80 percent of total cost subsidized from public revenues, the ratio of the pricing coefficients of the highest to the lowest categories of care is approximately 30:1.

Second, as the subsidy decreases, that is, as the proportion of total cost to be recovered from patient fees increases, the ratio of the high- to low-category pricing coefficients decreases markedly. Going from an 80 to 40 and then to 0 percent subsidy, the ratio falls from 30:1 to 12:1 and finally to 4:1. Thus, as the subsidy is reduced the latitude for cross-subsidization

is also reduced, with the lowest elasticity category taking on an increasingly greater burden of the cost of services. Even at a zero subsidy, however, the optimal pricing rules generate some cross-subsidization, with the pricing coefficient remaining less than one for the lowest service category but rising to above one for the highest categories.

Third, the pricing coefficients are moderately sensitive to the income elasticities. A lower elasticity for a given service produces a lower pricing coefficient on that service and a higher coefficient on other services. At a 40 percent subsidy and with income elasticities of 0, 0.3, and 1.5 for the three categories of services, the pricing coefficients are 0.12, 0.68, and 1.42, respectively. If the income elasticity for the highest category falls to 1.0, the pricing coefficients for the two lowest categories rise to 0.15 and 0.75, whereas the coefficient for the highest category falls to 1.20. This sensitivity illustrates the importance of further empirical studies to derive the demand characteristics of hospital services.

Fourth, the pricing coefficients are not highly sensitive to the income distribution through a range of realistic values from recent household surveys, but the coefficients are sensitive to the distributional objectives guiding government pricing policy. A standard normative principle in economics is that added income brings about increased welfare, but successive additions to income bring ever-smaller increments in welfare (this is the principle of diminishing marginal utility). The simulations are based on the subjective assumption that a given percentage increase in income is accompanied by an equal proportional decrease in the increment to welfare. The effect of this assumption is sufficiently humanitarian to lead to relatively aggressive redistribution objectives. If, instead, government planners believe that the increment to welfare falls faster (more slowly) than this assumption indicates, the extent of cross-subsidization implied by the pricing coefficients increases (decreases) markedly.

The actual pricing schemes in selected public hospitals in which there is an attempt to mount more than a nominal cost-recovery program are not greatly at variance with the implications of these simulations. An examination of the pricing coefficients produced by the simulations and the coefficients derived from the actual costs and prices for a level II hospital in Indonesia serves to illustrate their similarity. Expressly stated goals of Indonesia are to provide accessible hospital services for groups with the lowest incomes and to institute a program of modest cost recovery. On average, public hospitals in Indonesia subsidized about 80 percent of total cost, recovering the remaining 20 percent from user charges. With an 80 percent subsidy the optimal-pricing coefficients from the simulations are 0.03, 0.18, and 0.60, respectively, for the low-, middle-, and high-amenity (low-, middle-, and high-income elasticity) services. Based on the 1984 pricing guidelines of the Ministry of Health

as implemented in a small level II hospital, the implicit price coefficients for low-, middle-, and high-amenity services are 0.04, 0.30, and 0.51.[5] Thus, the Indonesian coefficients are taken as consistent with the optimum pricing simulation, and the pricing scheme used in the hospital is in line with the relatively egalitarian equity goals as well as a revenue objective of 20 percent cost recovery.

Risk Sharing

The high cost of hospital services, coupled with the randomness of many health needs, is the primary reason for the importance of insurance as a means of financing health services. The introduction of optimal pricing as a means of achieving the equitable distribution of care will not fully adjust for all the equity concerns that arise in the use of hospital services. If the total revenues raised through fees are to offset a substantial part of total costs, then the prices charged for all goods, even those demanded by low-income groups, must be a high proportion of unit costs. Hospital services are by far the most expensive health goods consumed and, depending on the illness, the cost of a hospitalization can easily amount to a multiple of per capita annual income. Modest fees for primary health care services at and below the level of the health center may be absorbed by the majority of the population without a risk of great financial loss, but the introduction of substantial fees for hospital services adds a risk of heavy financial costs to households and creates a need for insurance.

Some health needs occur randomly throughout the life of all individuals and thus, within a narrow confidence band, are predictable. But savings arrangements or contingent asset and credit markets may be inadequate to finance these predictable costs. In addition, other, unusual, health needs are unpredictable from the point of view of the individual. The cost of adequate treatment for many unpredictable illnesses can easily prove to be a catastrophic burden substantially affecting the welfare of the household. Health insurance improves efficiency by providing a form of earmarked savings for predictable risks and can also improve equity by spreading the risk of the cost of unpredictable illness among all households.

The availability of health insurance varies among developing countries. In those in which health services are heavily subsidized, government is implicitly covering individual risk, though this coverage is not actuarially based. Interpreted in this light, government provision of free services is a form of social insurance with no deductible and with zero coinsurance rates. Such subsidization limits the demand for more explicit forms of health insurance. High administrative costs and a lack of an institutional mechanism for collection in rural areas may also impede the use of health insurance. Also, an actuarially adequate premium may

be beyond the capacity of many households. For these reasons private insurance markets that cover individuals are often not well developed, and those that do exist usually cover only a small fraction of the population. Informal insurance arrangements, in which the financial risk is shared among the members of a community or extended family, may exist in some areas, but such mechanisms are apt to function unevenly. Employer plans providing either direct services or third-party risk coverage are also not common and cover only selected parts of the population. Government social insurance programs with a medical care component are more prevalent in Latin America but are rare and generally cover only a small proportion of the population in most African and Asian countries.

There is growing interest in government-provided insurance in many emerging economies, but experience in Latin America and the industrial countries demonstrates that adoption of social insurance must be done circumspectly if the programs are not to have unintentionally adverse effects. In many low-income countries the best choice may not be an explicit program of government-sponsored health insurance. In spite of the theoretical advantages of health insurance, specific schemes must be formulated carefully if substantial positive equity and efficiency benefits are to be realized.

Poorly devised health insurance schemes and those designed to benefit only specific population subgroups can result in a deterioration rather than an improvement in social welfare. For example, the use of health insurance to cover subsets of the population, such as civil servants or urban workers, in a partially monetized economy raises important equity issues. This is especially true if, as is generally the case, government revenues partially subsidize the superior quality and greater per capita quantity of services consumed by the insured population. Schemes can also create incentives to use resources inefficiently. Finally, the use of insurance introduces problems of administrative feasibility that are even greater than those brought about by the use of fees themselves because of the need for more careful accounting and administration, both in the collection of premiums and in the delivery of services. This difficulty in administration has been an important reason that insurance schemes have not been used more widely in poor economies. In the next chapter, we review experience with a selection of schemes and note potentially adverse effects in equity and efficiency. In the next few pages, we briefly review the principles of some risk-sharing alternatives and note salient aspects related to hospitals.

The type of insurance plan and efficiency. There are several possible ways to design health insurance schemes, and within the general formulation of each scheme there are many parameters that must be set correctly if

the scheme is to have a positive effect on the overall welfare of the population (Akin 1987). Broadly, we can classify insurance schemes into two principal types: third-party retrospective reimbursement and prepaid health care organizations. Among the many parameters defining the scheme, we focus our attention on the services to be covered, the magnitude of the health event to be covered, the population groups included, and the size of the premium and copayment. The choice of scheme and parameters closely affects the functioning of hospitals. Critical to an appropriate insurance plan are the implications it contains for client and provider behavior. Different reimbursement arrangements can bring about marked differences in the kind and quantity of services demanded by patients and given by hospitals.

In plans providing retrospective reimbursement on a fee-for-service basis, the individual pays a periodic premium to cover possible expenditures for specified services in the future. The premiums from different individuals are pooled and only used to cover services as needed at random by members. Hospitals are reimbursed retrospectively for the cost of each service provided. This payment system gives hospitals an incentive to add to revenues by maximizing the volume of services provided per patient and providing the most costly services possible. As is true of all insurance systems, patients have no incentive to be concerned with the cost of care unless there are cost-sharing provisions. The effect of this hospital reimbursement system is to increase health care costs. One possible means of countering these adverse incentives is a regulatory process whereby hospitals can be denied reimbursement for services determined to be unnecessary. Of course, such a process is itself costly.

An alternative retrospective reimbursement plan reimburses the hospital a fixed amount per case or admission, which is to cover all services provided to the patient during his or her stay. The cost-containment incentives are superior to those of fee-for-service reimbursement, as the hospital is encouraged to minimize service inputs per admission. On the other hand, the hospital may try to maximize the number of admissions for which its margin of reimbursement above treatment costs is greatest (that is, hospitals may try to choose healthier patients). To mitigate this problem, diagnostic and other patient or procedural information may be used to group patients into categories, each with its own case payment rate (Fetter and others 1980). The difficulties of defining diagnostic categories, establishing and periodically adjusting case reimbursement rates, and policing and administering this type of plan can be formidable, however. Case-based reimbursement raises concerns about quality of care, given the incentives to minimize inputs per case. These concerns can be addressed through some combination of quality-based competition among hospitals or insurers, and utilization review and other forms of regulation.

Finally, prepaid capitation provider plans (an alternative terminology is health maintenance organization, or HMO) remove provider incentives to increase the cost of care. In these plans, the provider is prepaid a fixed amount to cover health care needs during a specified time interval and then delivers services as needed without further reimbursement. This arrangement focuses on the population rather than on the providers to be reimbursed; therefore, a capitation health plan should involve all levels of personal health services so that the care of individual patients may be managed cost-effectively. One form of this model is direct insurance, in which the provider and insurer are the same institution and thus respond to the same incentives. In contrast to the retrospective reimbursement plans, the capitation plan introduces provider incentives to reduce both the total number of admissions and quantity of services provided per case. With these incentives providers may lower the quality of care in order to reduce the cost of services. As in case-based reimbursement, the means of mitigating such effects would be either competition among prepaid plans or some form of regulation that would ensure quality.

In all these schemes there is an incentive, termed "adverse selection," for high-risk patients, whose health needs will probably exceed the cost of the average claim, to join and the converse incentive for low-risk people. There is also an incentive, termed "selection bias," for insurers to exclude high-risk persons from their risk pools in an attempt to maximize the margin between premium income and claims paid. Adverse selection can greatly increase the cost of an insurance program if there are significant differences in relative risk across insurance groups, and selection bias can create equity problems if high-risk persons are unable to obtain insurance. The effects of adverse selection and selection bias can be reduced by requiring enrollment in insurance plans across broadly defined client groups or by organizing insurance plans around other broadly defined groups.

Insurance can also create an incentive for clients to change their behavior, which can affect the cost and quantity of services demanded. Such changes occur because the clients are less concerned about possible financial loss. For example, high-risk pregnancies and the use of hospital services for delivery might increase if the cost of obstetrics is covered by insurance. This phenomenon is called "moral hazard." The effect of moral hazard can be reduced by limiting benefits or by requiring a copayment or deductible from clients for part of the cost of services. The argument for including cost sharing as part of an insurance scheme is similar to that for instituting user charges in formerly free care systems. Excess use of services resulting from moral hazard is equivalent to what is often termed "unnecessary utilization," which arises in free care systems. In both cases, the high demand is the rational response of consumers faced with a very low priced good.

Equity and government insurance. Government-sponsored health insurance for specific employment groups such as civil servants has been proposed as a means of transferring resources from urban to rural areas and from high-income to low-income health service users. The argument is that most high-cost hospital services are consumed in urban areas and by specific employment groups whose income is higher than average, and that the revenues obtained from insurance premiums and copayments in urban areas can be used to cross-subsidize rural services for poorer people. This is a valid argument as long as the revenues obtained are greater than the cost of services provided; in this case the cross-subsidization will go in the direction intended. Review of actual experience, however, reveals that revenues seldom exceed the cost of services. Instead, the government subsidy of urban hospital services covered by insurance is often substantial and, depending on the source of revenues, the equity effects are adverse. For example, the ASKES insurance program provides free hospital care for Indonesian civil servants but reimburses hospitals only about 15–25 percent of the cost of an average inpatient stay. Furthermore, the hospitalization rate for ASKES beneficiaries is about five times the national average (Prescott 1991). Thus, moral hazard magnifies the negative impact on equity of providing insurance coverage for subsidized hospital services to a part of the population that is relatively well off. Government health insurance plans, therefore, need to be established in conjunction with a careful choice of services covered, a knowledge of unit costs, and premium and copayment levels that will achieve the desired distributional objectives. One possible way to reduce the adverse equity problem is for the government-sponsored plan to include only a defined minimum benefit package covering amenity levels and services that provide adequate care for the average consumer, to charge fees for services excluded by the benefits package, and to leave the provision of additional coverage to private insurance.

Prepaid plans and hospitals. Prospective payment plans can be organized on a modest scale and provide a means of increasing revenues for specific hospitals or groups of hospitals within a community. Through use of a prepayment plan, rural communities served by district hospitals can achieve greater financial autonomy and reduced dependence on central budget sources. Such plans can also yield improved efficiency in the use of providers if the prepayment amount covers all personal health services for plan members, not just hospitalization. In this case, providers have the incentive to steer enrollees in the plan to the least costly service delivery setting.

The advantages of a prepaid plan are that they reduce the incentives for excessive consultations, excessive diagnostic tests, and higher drug use and surgical rates that are reported to exist with a retrospective

payment system (Shimmura 1988). The administrative costs of prepaid plans are less than those of reimbursement schemes because of the reduced need for billing information and records, and the requirements by management for information within the hospital are less than those in retrospective payment schemes. As is the case for retrospective plans, however, a copayment may be needed to reduce demand for less necessary treatment and to improve economic efficiency.

Type of service covered. There is a conflict between the need to cover only catastrophic losses to reduce the cost of an insurance plan and the need to include broader, noncatastrophic coverage to avoid the introduction of spurious overuse of higher-cost hospital services. This conflict exists under both retrospective and prospective plans, although it is most relevant for retrospective plans because coverage under such plans is often narrow, whereas prepaid plans commonly provide broad coverage of services. Some authors (for example, Griffin 1989) have recommended that only catastrophic costs be covered. The difficulty lies in defining the basis and limits of the coverage. Perhaps the best possibility would be to set the limits of the coverage on the basis of individual annual health expenditures or (slightly less desirable) on the basis of an individual health event. Insurance would become applicable only above a certain absolute amount. This arrangement would place a burden on the consumer of keeping records and then filing claims as justified. The requirements for literacy, numeracy, and organization may be too great, however, for low-income groups in many countries. A second possibility would be to define "catastrophic" in terms of specific services to be covered, such as inpatient care of more than a certain number of days, emergency services, or specific diseases. This arrangement, although more practical in terms of recordkeeping and administration, could greatly distort the use of hospital services by encouraging providers to use unwarranted services in order to claim insurance coverage. Controlling such misuse of services would probably require regulation and monitoring.

Limiting insurance to catastrophic coverage also results in a loss of the efficiency gain provided by the credit or savings function that comes from covering the cost of subcatastrophic random care. By extending coverage to selected services provided by health centers and hospital outpatient departments, copayment levels for the lower referral levels can be set so that patients are encouraged to use these less costly settings. The cost of the services can then be offset from the insurance fund composed of accumulated premiums.

Summary

In this chapter we have outlined the principal cost-recovery alternatives to the financing of hospitals from public sector budgets. All the alternatives have potential effects on the use and provision of services that are both beneficial and detrimental to welfare. The planner's dilemma is to weigh the benefits and costs associated with the alternatives in order to design a scheme that is most suitable for a given environment. Many of the effects can be attributed to the incentives created by the financing mechanism for consumers and providers. Responses to these incentives may differ across institutional and cultural environments. For example, the behavior of providers depends on whether they are motivated by quantity, revenue, or profit maximization; cost minimization; quality objectives; or some combination of these. The behavior of consumers is affected by their education, income, and perception of the quality of facilities; the range of providers from which to choose; and the degree of practical control the consumer has over treatment.

There are dozens of variations on the basic fee and risk-sharing alternatives. Table 4-3 provides a very general summary of the incentives inherent in six broad financing options, ranging from high fees and no insurance through high fees and partial insurance to capitation payments for managed health care. Some of the options, such as high user charges and no insurance, clearly create incentives that adversely affect equity. Other options, such as free public provision of services or, equivalently, full insurance coverage without cost sharing, create incentives that encourage inefficient use of resources. Practical systems that avoid or control these adverse effects through exemptions, partial coverage, capitation, or price discrimination are apt to vary and be tailored to the situation in a given country or geographic area. In the next chapter, we examine the actual experiences of hospitals in recovering costs with the financing options outlined above.

Table 4-3. Summary of Incentives Inherent in Alternative Financing Policies

	Incentives	
Policy	*Providers*	*Consumers*
High user charges (fee-for-service)	Maximize billable services, possibly constrained by awareness that consumers' ability to pay is limited	Incentive to reduce use, but without insurance, access to expensive services is difficult for many
Fee-for-service plus third-party insurance with no cost sharing	Maximize billable services; producer is virtually unconstrained by consumers' ability to pay	For the insured population, quantity demanded is the same as when money price is zero; no cost consciousness
Third party fee-for-service insurance with cost sharing	Maximize billable services	Some cost consciousness, but demand still greater than if consumers faced full prices
Insurance coverage using prospectively determined reimbursement rates per case	Minimize costs per patient, maximize admissions of patients whose treatment costs are less than reimbursement amount (that is, shift case mix); report diagnoses or procedures for ambiguous cases that maximize reimbursement	Extent of cost consciousness depends on existence and magnitude of cost-sharing provisions
High fees plus insurance coverage of only part of the population	Maximize billable services; focus on the insured from whom reimbursement is guaranteed	Inequitable access to care because price to the insured is much less than to the uninsured
Insurance coverage financed through capitation payments and administered as a managed care system; also direct insurance	Keep total health care costs for covered population below sum of capitation payments; this may lead to cost-effective allocation of resources or underprovision (especially if there is no competition among providers)	Guided through the network of providers; typically faced with high money prices if they violate the structure of the managed care system

Appendix 4A. Optimal Prices for Hospital Services

The crux of a cost-recovery policy lies in setting prices and insurance parameters that promote the equity, efficiency, and revenue goals that are the objectives of financing. Fortunately, the general problem of setting prices in the public sector has received growing attention by welfare theorists during the last twenty years, and a flexible theory of public sector enterprise pricing has arisen. The literature is capable of illuminating important pricing and insurance issues and, in particular, holds promise for application to hospitals.

Fees

A well-known result of welfare economics states that to achieve optimal (Paretian) efficiency, prices for all firms should be set equal to marginal cost:

$$P = MC. \tag{4-1}$$

Were it not for market failure, and equity, the setting of fees for public hospitals could follow this simple rule to achieve economic efficiency. However, market failure and a concern for equity limit the application of marginal cost pricing for hospitals (Baumol and Bradford 1970). Optimal pricing principles for public enterprises provide an adjustment to marginal cost pricing that can be used to address questions of market failure and equity.

Public enterprise pricing (Ramsey pricing). To achieve social and political objectives, hospital services are provided at less than cost and supported by government subsidies. Additionally, some hospitals may have declining costs throughout the relevant range of production, and marginal cost will be below average cost. Thus, a price equal to marginal cost will *perforce* entail a loss, and a subsidy will be required. A pricing rule developed by Ramsey (1927) and later modified by Boiteux (1971) maximizes welfare constrained by a budget equal to cost-recovery revenue plus a fixed subsidy. This rule, which forms the basis for pricing in modern public enterprises, can be stated in terms of the price-cost margin (that is, the ratio between the price-cost difference and price) as

$$\frac{P_i - MC_i}{P_i} = \left[\frac{\lambda - 1}{\lambda}\right]\frac{1}{-E_i} \quad i = 1 \ldots n \tag{4-2}$$

where E_i is the elasticity of demand for service i and λ is the added benefit or shadow value of an additional unit of budget (subsidy).

Although λ is not directly observed, the Ramsey rule nevertheless provides important implications for pricing (elaborated in Bos 1985). The rule states that the price-cost margin is inversely proportional to the own price elasticity of the good in question. The Ramsey rule does not provide for explicit cross-subsidization, because the deficit is spread over all services whether luxuries or not, but the rule does affect distributional objectives through the role that price elasticities play in setting prices. In effect, the inverse elasticity of each good is multiplied by the same value, $(\lambda - 1)/\lambda$, to obtain multipliers that can be used to mark down (or mark up) costs so that the budget deficit (or surplus) is distributed over all goods. Lower (absolute value) elasticities will have higher absolute price-cost margins. If the goods to be priced are produced at a loss (that is, the price-cost margin is negative, as is generally the case for public hospitals), then the rule leads to relatively lower prices of price inelastic goods and higher prices of elastic goods. If the goods purchased by lower-income households are price inelastic, the equity effects of the pricing rule will be favorable compared with, say, those of marginal cost pricing.[6]

The Ramsey rule is capable of practical application. Given a predetermined level of subsidy, costs and estimates of the price elasticities for n goods, the values of the prices, and the critical proportion can be imputed from the n sets of Ramsey relationships and the budget constraint. In practice all the elasticities required may not be available from econometric estimates, but subjective estimates together with the restrictions of demand theory may be sufficient to identify the required prices.

Public enterprise pricing with distributional weights (Feldstein pricing). The equity benefits of applying the Ramsey rule to deficit budgets are not the result of directly including distributional considerations in the planning process but are an artifact, or side aspect, of the Ramsey result. The equity effects of Ramsey pricing, although a move in the right direction, fall short of the optimal effects that can result from directly including distribution. Recent literature that incorporates distributional goals as well as efficiency as a policy objective has grown as a result of seminal work by Feldstein (1972a, 1972b, 1972c). The Feldstein method derives a rule for pricing that maximizes social welfare defined as a weighted function of the consumption of individual households with varying incomes and consumption patterns affected by the level of their incomes. The weights explicitly recognize that the value of additional consumption decreases with rising household income. The Feldstein rule can be expressed as

$$\frac{P_i - MC_i}{P} = \left[\frac{\lambda - R_i}{\lambda}\right]\frac{1}{-E_i} \quad i = 1 \ldots n \tag{4-3}$$

where R_i is the "distributional characteristic" of the good and is, in fact, the weighted function of household consumption. In general the specification of R can be complex. Feldstein suggests an approximation that allows the calculation of the R_i in terms of the mean and variance of income, an estimate of the income elasticity of demand and a normative value for the elasticity of social marginal utility with respect to income.[7]

The implications of the Feldstein rule for pricing can be elaborated. Lower values of R yield higher prices.[8] The value of R_i is inversely related to the income elasticity of demand (see Sherman 1989). Thus, goods with higher income elasticity of demand have relatively lower values of R and higher optimal prices.

Ramsey prices can be used for reference. When the income elasticity of demand is 1, R_i will equal 1. Equation 4-3 then reverts to the Ramsey rule. If the income elasticity of demand exceeds 1, the value of R will be less than 1 and the Feldstein price will exceed the Ramsey price. Conversely, if the income elasticity is less than 1, the Feldstein price will be less than the Ramsey price. Thus, the Feldstein rule increases the prices paid for services consumed by the rich and reduces the prices of services paid by the poor. Cross-subsidization will occur if the price elasticities of the goods consumed by the rich are sufficiently high (in absolute value) that $P > MC$.

The Feldstein prices will also vary with the distribution of income and the rate at which marginal social valuation of income diminishes. The greater the inequality of income distribution or rate at which the marginal valuation of income falls off as income rises, the lower will be the value of R for any given income elasticity. Using the Feldstein approximation for R_i (given in note 7) and an estimated variance of household expenditure for Côte d'Ivoire (Glewwe 1988), and choosing an elasticity of social marginal utility of –0.5, we find that the value of R_i is 1.3 for goods with an income elasticity of 0, 1.1 for goods with an income elasticity of 0.5, and 0.8 for goods with an income elasticity of 1.5.

Two-part pricing. Optimal public enterprise rules have also been derived for two-part pricing, in which the first part is a flat fee for all consumers to have access to services and the second part is a price for services used. Optimal two-part pricing rules are relevant, for example, if an admission fee is charged for hospital inpatient services or an outpatient ward in addition to fees for specific services received. In general two-part pricing rules with distributional weights are similar to the Feldstein rule set out above. The principal difference is that fixed-fee receipts are added to the subsidy in the budget constraint with the result that the value of λ changes. We do not pursue two-part pricing further here because it would carry the discussion too far afield, but we do note that the problem is tractable (Feldstein 1972b; Sherman 1989).

A suggested pricing procedure. In application, a hospital has too many services to allow the practical identification of all the information that would be required to implement Ramsey or Feldstein pricing fully. Instead, a procedure is suggested that would approach, but not fully achieve, the optimal pricing solution and promote increased efficiency and equity in comparison with either zero prices or arbitrary pricing not directly relating costs and demand.

The procedure is as follows:

- Classify hospital services into large groups of bundled services and differentiate the groups by subjective estimates of the income elasticity of demand into, say, three broad categories: low-income elasticity goods (including inferior goods), normal goods, and luxury goods.
- Carry out a unit cost analysis using, say, step down disaggregation of cost as discussed in the previous chapter. Calculate the unit cost (c_i) of the services to be priced and the estimated total recurrent cost of the principal service groups. A second subscript can be added to unit cost to indicate the low-, normal-, and high-income elasticity categories as c_{iL}, c_{iN}, and c_{iH}, respectively. The total recurrent costs of the groups can be designated C_L, C_N, and C_H, where, for instance $C_L = \Sigma c_{iL} Q_{iL}$. Total recurrent costs are then

$$C = C_L + C_N + C_H .$$

- Identify the magnitude of the hospital revenue objective as the expected total recurrent costs (C) less the government subsidy (S),

$$D = C - S .$$

Then, the problem to be addressed in the next steps (d and e) is the selection of markdown coefficients (δ_i) for unit costs, so that the prices determined,

$$P_i = \delta_i c_i ,$$

will completely distribute the deficit among the categories, that is,

$$\Sigma \delta_{iL} c_{iL} Q_{iL} + \Sigma \delta_{iN} c_{iN} Q_{iN} + \Sigma \delta_{iH} c_{iH} Q_{iH} = D$$

or, if δ_i are the same within each group,

$$\delta_L \Sigma c_{iL} Q_{iL} + \delta_N \Sigma c_{iN} Q_{iN} + \delta_H \Sigma c_{iH} Q_{iH} = D ,$$

which can be written

$$\delta_L C_L + \delta_N C_N + \delta_H C_H = D. \tag{4-4}$$

- Determine the distribution parameters, R_i. If information on income distribution and income elasticities is available, R_i can be calculated for a chosen elasticity of marginal social utility of income. If information is lacking, a practical approximation is to set the value of R_H at 0.75 for the high-income elastic group, R_N at 1 for the normal group, and R_L at 1.25 for the low-income elastic group of services.
- Finally, choose markdown coefficients that are consistent with Feldstein prices for each group. This is done through application of equation 4-3. In the absence of better information, use average unit costs (c_i) as a substitute for marginal costs (MC_i) in equation 4-3. Assuming that price elasticities are comparable within the groups, equation 4-3 yields expressions for the markdown coefficients,[9]

$$\delta_i = \frac{\lambda E_i}{\lambda - 1 + \lambda E_i} \qquad i = L, N, H \tag{4-5}$$

The three equations represented by 4-5, together with equation 4-4, can then be solved simultaneously for λ, δ_L, δ_N, and δ_H, which can in turn be used to set prices within the groups.

As an example we take a hospital with total costs of $C = 250$, which can be broken down into three primary service categories with group total costs of $C_L = 100$, $C_N = 100$, and $C_H = 50$. (To provide a frame of reference, the three service categories can be thought of as a bundle of services associated with inpatient care in wards, semiprivate care with moderate amenities, and private room care with substantial amenities.) The subsidy is 100, giving a deficit of $D = 150$ to be financed from revenues on the sale of services. Using $R_L = 1.25$, $R_N = 1.00$, and $R_H = 0.75$ and assuming the *price* elasticities are –0.05, –0.3, and –1.00 for the income-elastic categories of low, normal, and high, respectively, one can solve the budget constraint (equation 4-4) and the markdown coefficient (equation 4-5) to obtain a markdown coefficient of $\delta_L = 0.12$ for low-income elastic services, $\delta_N = 0.79$ for the middle category of services, and $\delta_H = 1.21$ for high-income elastic services. The burden of the deficit then falls on the high-income elastic category, which generates net profits, whereas the lower-income elastic category receives a large share of the cross-subsidy profits and the government subsidy.

Sensitivity of the pricing coefficients to the subsidy. The government subsidy together with profits is the primary determinant of the markdown coefficients. When the subsidy is very low the proportion of cost recovered must be high even for the low-income elasticity good. In contrast, when the subsidy is very high none of the goods is sold at a profit, and the markdown on the low elasticity good becomes nearly complete. The

data in table 4A-1 illustrate how the markdown coefficients vary with the level of subsidy. The situation described is similar to the preceding example with the exception that the R_i are calculated from the parameters specified in the table. When the subsidy is 100 the high-elasticity category is sold at a considerable markup above costs ($\delta_H = 1.42$) and the lower two categories are cross-subsidized. If the subsidy should rise to 200 (80 percent of total costs), δ_L would fall to a negligible value (0.02) and even luxury services would be subsidized ($\delta_H = 0.59$). In contrast, if the subsidy should fall to 0, δ_L would rise to 0.4, δ_N would go above unity to 1.2, and δ_H would rise to 1.8 to provide additional profits to cover the services provided at a loss in the low-income elastic category.

Sensitivity of pricing coefficients to the distributional parameters. The pricing coefficients are also sensitive to the distributional characteristics of the services, that is, the parameters determining R_i. Data in tables 4A-2 through 4A-4 illustrate the change of the markdown coefficients with alternative values for the relative variance of the income distribution, the income elasticity of marginal social welfare (η), and the income elasticities of the services to be priced.

The markdown coefficients are not highly sensitive to the income distribution through a range of values based on recent household surveys in Ghana (Glewwe and Twum-Baah 1991) and Côte d'Ivoire (Glewwe 1988). Ghana represents a moderately equitable distribution of income with a normalized standard deviation of 0.70 (a variance ratio of 0.49), whereas Côte d'Ivoire represents a considerably less equitable distribution with a normalized standard deviation of 0.91 (a variance ratio of 0.84; in comparison, the variance ratio for the United States in 1970 was 0.55). The suggested distributional parameters to be used in the absence of specific information (see step [d] above) used a normalized standard deviation of 0.64 (a variance ratio of 0.80).

The income elasticity of marginal social welfare (η) is a subjective parameter. The more egalitarian the social objectives determining government policy the higher the value of η that should be chosen. A value of $\eta = -1$ implies that a 10 percent increase in income is associated with

Table 4A-1. Sensitivity of Pricing Coefficients to the Subsidy

Pricing coefficient	*Income elasticity*	*Own price elasticity*	*Subsidy (percent of total cost)*		
			0	*40*	*80*
δ_L	0.0	–0.05	0.41	0.12	0.02
δ_N	0.3	–0.40	1.18	0.68	0.18
δ_H	1.5	–1.00	1.80	1.42	0.59

Note: Eta = -1.0; total cost = 250 (cost of L = 100, cost of N = 100, cost of H = 50).

Table 4A-2. Sensitivity of Pricing Coefficients to the Relative Variance of the Income Distribution

			Relative variance of income		
Pricing coefficient	*Income elasticity*	*Own price elasticity*	*0.49 (Ghana)*	*0.65*	*0.84 (Côte d'Ivoire)*
δ_L	0.0	–0.05	0.13	0.11	0.09
δ_N	0.3	–0.40	0.68	0.64	0.59
δ_H	1.5	–1.00	1.42	1.48	1.63

Note: Eta = –1.0; total cost = 250 (cost of L = 100, cost of N = 100, cost of H = 50); subsidy = 100; net deficit = 150.

Table 4A-3. Sensitivity of Pricing Coefficients to the Income Elasticity of Marginal Social Welfare

			Elasticity of marginal social welfare (–Eta)				
Pricing coefficient	*Income elasticity*	*Own price elasticity*	*–0.1*	*–0.5*	*–1.0*	*–1.5*	*–2.0*
δ_L	0.0	–0.05	0.26	0.18	0.12	0.08	0.05
δ_N	0.3	–0.40	0.76	0.73	0.68	0.57	0.47
δ_H	1.5	–1.00	0.94	1.14	1.42	1.69	1.97

Note: Total cost = 250 (cost of L = 100, cost of N = 100, cost of H = 50).

Table 4A-4. Sensitivity of Pricing Coefficients to the Income Elasticity of Good 3

			Income elasticity of good 3		
Pricing coefficient	*Income elasticity*	*Own price elasticity*	*0.5*	*1.0*	*1.5*
δ_L	0.0	–0.05	0.17	0.15	0.12
δ_N	0.3	–0.40	0.82	0.75	0.68
δ_H	(see at right)	–1.00	1.00	1.20	1.42

Note: Eta = –1.0; total cost = 250 (cost of L = 100, cost of N = 100, cost of H = 50); subsidy = 100.

a 10 percent decrease in marginal social utility. The values in the table range from extremes of $\eta = -0.1$ to -2.0. Through this broad range the markdown coefficients vary considerably from no cross-subsidization from profits (on the high-elasticity services) at the lower value of η to considerable cross-subsidization from profits at the higher value of η. The suggested distributional parameters to be used in default of other information employ a unit elastic value for η.

Because the services are gathered broadly into only three groups the income elasticities are highly approximate. Also, knowledge of income elasticities depends on household survey and econometric information that may not be available or affordable. The most critical category of goods to classify are those with the higher-income elasticities that provide the profits for cross-subsidization. The sensitivity analysis suggests that the markdown coefficients on the categories that are less income elastic are not sensitive to a plausible range of elasticities for the luxury category of goods. The markdown coefficient for the high-income elastic good is only moderately sensitive, varying by less than 20 percent (from $\delta = 1.2$ to 1.4) as the elasticity goes from 1.0 to 1.5. The default values for R_i assume income elasticities of 0.0, 0.3, and 1.0.

Marginal compared with average costs. The above procedure assumes constant costs, that is, that marginal costs equal average costs or, as will be made clear below, that the ratios of marginal and average costs are the same for all service groups. Commonly, this is taken to be a reasonable approximation of reality. The evidence reviewed in the previous chapter was inconclusive for developing countries, but the consensus, after many econometric studies in industrialized countries, appears to reject hospitals as a declining cost industry. If additional information on marginal costs is available, however, and the marginal cost and average cost ratios are not the same, the costing procedure can be modified to set prices with greater accuracy (in the sense of achieving the welfare and efficiency objectives). If marginal and average costs are not equal the markdown equations become

$$\delta_i = \frac{\lambda E_i}{\lambda - 1 + \lambda E_i} \cdot \frac{MC_i}{c_i} \quad i = L, N, H. \tag{4-6}$$

Because the δ_i are determined simultaneously they will be affected in comparison with the prices set by using equation 4-5 only if the ratios MC_i/c_i differ across the groups.

The number of groups. The procedure outlined above is generalizable to more than three groups, say n. Additional categories only require the addition of further equations parallel to 4-5 and an appropriate modification to the budget constraint. Determination of the values of the markdown coefficients, δ_i, will then require the solution of $n + 1$ simul-

taneous equations, and this, as a practical matter, is easily executed using nonlinear iterative techniques. The purpose of the sparse choice of only three groups is to keep the problem tractable for small hospitals that do not have a well-established costing system. It is also probable that the additional welfare and efficiency diminishes rapidly with an expansion of the number of groups much beyond four or five, as long as the elasticity characteristics (both price and income) within the groups do not vary too greatly.

Health Insurance and Optimal Pricing (Harris pricing)

The introduction of insurance alters the pricing rules set out earlier for an environment without insurance. Harris (1979) has extended the Ramsey analysis to derive optimal prices for a hospital operating in an environment in which all members of the community are covered by insurance. The addition of insurance means that the optimal price will vary with the copayment rate, the price elasticity of demand, and the marginal utility of income. The Harris analysis is explicitly designed for retrospective reimbursement insurance that has a fixed premium equal to the policy's actuarial value and fixed coinsurance rates. The full derivation of Harris's results would take the discussion too far afield, but the results can be summarized in the following pricing rule:[10]

$$\frac{P_i - MC_i}{P_i} = \left[\frac{\lambda - 1}{\lambda}\right]\frac{1}{-E_i} + \frac{I_i}{\lambda}, \tag{4-7}$$

where

$$I_i = -\frac{r_i}{x_i} cov\,[\,\mu_s\,, x_{is}\,] + (1 - r_i)$$

and, in addition to the notation introduced in the previous section, r_i is the coinsurance rate (determined exogenously) for service i, x_i is the demand of the average consumer for service i, and cov $[\mu_s\,, x_{is}]$ is the covariance of the marginal utility of income (μ_s) and the consumption of service i across states of health, s.

This rule differs from the Ramsey optimal price (equation 4-3) with the addition of I_i/λ, which adjusts for risk effects across health states. The rule states that the optimal price will be increased by a decrease in the copayment, r_i, or the covariance term, cov $[\mu_s, x_{is}]$. The covariance term describes the extent to which the demand for specific health services is associated with costly health states and large sacrifices in consumer welfare. If the health service is covered by insurance, the covariance between the marginal utility of income and the quantity of service consumed will be reduced and, because the sign on the covariance term in equation 4-6 is negative, the optimal price is increased. The effect of the copayment rate can be regarded as a correction for the efficiency

distortion of subsidized services provided under insurance. The covariance term represents the equity benefits derived from spreading the risk of expensive health states across the pool of insured persons.

Harris used this method to reexamine the prices for services in a United States hospital. Sufficient data were available to allow the pricing of a diverse set of surgical, diagnostic, intensive care, and ward services. To be applied in most hospitals in developing countries, the analysis would require fewer and more aggregated pricing categories. The information required for the analysis would be estimates of the unit cost of services, the price elasticities of demand, and a survey of the economic and health status of consuming households.

Harris pricing is applicable when most or all of the population is covered by insurance; any of the population not covered would be served either by subsidized prices or at full prices as determined by income screening. Thus, Harris pricing may be appropriate for a comprehensive system of national insurance or for smaller, circumscribed communities served by a given health plan. If insurance covers only a small proportion of the population, however, the Harris prices need to be modified.

Adding distributional weights. There are several possible extensions of Harris prices that have not yet been discussed in the literature on optimal pricing but that could be useful for setting policy and should be addressed by future research. First, distributional objectives could be introduced directly into the optimization objectives. The distributional effects of Harris pricing are attributable to the welfare gain from spreading the risk of exceptionally costly illness across the pool of insured households. As with Ramsey prices, the inequity of the distribution of income is not explicitly recognized, but the protection insurance provides against market prices at least partly mitigates this failure. Distributional objectives parallel to those specified by Feldstein could be combined with Harris's specification to obtain a more complex rule, introducing the distribution term, to set prices.

Optimal coinsurance rates (Arrow prices). As a second modification to Harris prices, it would be possible to reverse the optimization procedure and derive the optimal coinsurance rates, given specified prices. Harris prices are optimal, given a prespecified set of coinsurance rates. From the point of view of a single consumer, Arrow (1976) discussed the welfare implications of changes in coinsurance rates in the situation where prices for health services are determined by the market but did not actually derive optimal rates or consider the distributional implications. His findings could, however, be modified to encompass nonmarket prices and to yield optimal coinsurance rates. Optimal coin-

surance rates, given prices, are potentially useful when a large part of the population is not served by insurance.

A two-step procedure to set prices could provide substantial welfare gain in situations in which the insured comprised only a minority of the population. Feldstein pricing, which does incorporate the distribution of income, could be used to set prices in the first step. The Arrow technique would then be used to determine the optimal coinsurance rates given the prices set in the first step.

Notes

1. For nontechnical reviews of the issues, see Griffin 1987; Hoare and Mills 1986; de Ferranti 1985; and Akin, Birdsall, and de Ferranti 1987. A more technical, yet accessible, discussion of pricing in the social sectors in general is in Jimenez 1987. Finally, excellent technical reviews of public sector pricing alternatives are Atkinson and Stiglitz 1980; Bos 1986; and Sherman 1989. This chapter draws substantially on all these studies.

2. A general discussion of incomplete markets in given in Laffont 1989.

3. This phenomenon can be made more explicit if we set out the algebraic relation between the net government subsidy and fee retention and examine the change in the net subsidy with the percentage of fees retained in order to identify an optimal retention policy. Given the actual level of autonomous revenues received (H_A) and a proportion (ρ) of these revenues returned to the central treasury, the net government subsidy (S) can be defined as

$$(1) \qquad S = S^* - \rho \cdot H_A \qquad 0 < \rho < 1 ,$$

where S^* is the original budgeted subsidy. The relation between the actual autonomous revenues (H_A) and potential autonomous revenues (H_P) can be written as

$$(2) \qquad H_A = r \cdot H_P \qquad 0 < r < 1 ,$$

where r, the proportion of potential revenues collected, is a measure of the efficiency of cost recovery. Finally, the behavior of collecting agents is summarized by

$$(3) \qquad R = f(\rho) \qquad \partial f / \partial \rho < 0 ,$$

which states that the efficiency of cost recovery is inversely related to the proportion of fees returned to the central treasury. Substituting (2) and (3) into equation (1) and setting the derivative of S with respect to ρ equal to zero allows a solution for the value of ρ that minimizes the net subsidy:

$$(4) \qquad \rho = -f / (\partial f / \partial \rho) .$$

If, to take a simple example, $r = 1 - \rho$, then the optimal fee retention policy would be to set $\rho = 0.5$. In practice, the relation between ρ and r described by

equation (3) has not been empirically determined and will have to be found by incremental policy experiments.

4. With some liberal modifications, these are based on Ellis 1987.

5. The implementation of the pricing guidelines varies considerably across provinces and districts in Indonesia. As implemented in Mataram hospital, the 1984 basic inpatient charge for beds was Rp. 600 for class III, Rp. 2,500 for class II, and Rp. 6,000 for class I amenity levels. Drugs and diagnostics are included in the class III price but excluded in classes II and I. The addition of an estimated charge for drugs and diagnostics of Rp. 2,500 per bed-day (about 15 percent of average costs) brings the total estimated price per bed-day for the three amenity classes to Rps. 600, 5,000, and 8,500. The estimated average cost per bed-day (1984 prices) is Rp. 16,500. Dividing the bed-day charges by the cost gives implied pricing coefficients of 0.04, 0.30, and 0.51.

6. If, however, the goods are produced at a profit (which is seldom the case for hospitals), then the converse result holds—the prices of price inelastic goods will be relatively higher and the distributional effects will be adverse.

7. Using a log-normal distribution of income and a constant elasticity of marginal social utility of income and normalizing by the mean of income, we can estimate R_i by

$$R_i = \left\{(1 + V)^{(1/2)\eta(1+\eta)}\right\} \cdot \{1 + V\}^{-\eta\alpha_i},$$

where η is the elasticity of marginal social utility of income, α_i is the income elasticity of good i, and V is the relative variance of income, $\sigma_y^2/\bar{y}^2$. The elasticity of marginal social utility of income provides an intuitively natural characterization of equity (Feldstein 1972a, p. 36).

8. When $R = 0$, equation 4-3 becomes

$$\frac{P_i - MC_i}{P_i} = \frac{1}{-E_i},$$

which is the expression for the price that would be charged by a profit-maximizing monopolist.

9. If we substitute $P_i = \delta_i c_i$ for price, equation 4-3 can be written,

$$\frac{\delta_i c_i - c_i}{\delta_i c_i} = \frac{\lambda - R_i}{\lambda} \cdot \frac{1}{-E_i}$$

or, canceling out the unit costs,

$$\frac{\delta_i - 1}{\delta_i} = \frac{\lambda - R_i}{\lambda} \cdot \frac{1}{-E_i}$$

which can be solved for δ_i as above.

10. The equation given incorporates several assumptions made by Harris to simplify the application of the pricing rule. The most important of these assumptions is to define the marginal utility of income (λ_s) so that there is constant absolute risk aversion at all levels of income.

5. *Hospital Financing Experience*

During the 1980s, many countries implemented user fees in public hospitals, increased existing fees, changed policies with regard to the dispensation of fee revenues, or implemented or modified social insurance schemes to pay for hospital services. The principal forces behind these changes were a reduction in the availability of recurrent resources from central or local governments (for example, in Indonesia; Prescott 1991) and rapid escalation of hospital costs (for example in Brazil; McGreevey 1988). Many governments perceived a need either to increase the availability of resources to health facilities or to reduce the extent to which they subsidized these facilities.

As discussed in chapter 4, the salient features of alternative financing programs can be broadly categorized as user fees and insurance, although there are differences in how policies are implemented in specific countries. The great variety of policy choices—such as exemptions, price discrimination, differing subsidies for fees and premiums, deductibles, coinsurance rates, excluded services, and reimbursement mechanisms for risk-sharing schemes—give planners considerable freedom to tailor a health financing system to planning objectives. The actual experience of countries can illuminate the possible implications of policy alternatives.

In this chapter we review user fee and insurance programs that have been implemented in various countries during recent years. The focus of the review is on the three objectives of alternative hospital financing policies: revenue generation, efficiency, and equity. A review of experience provides some important practical lessons for the relation of financing mechanisms to the achievement of these policy goals. We consider the institutional features of user fee and insurance programs and elaborate on how alternative financing policies have interacted with problems of service distortion, adverse technology choices, and inappropriate referral. First, we present data on the revenue performance of fee and

insurance programs in several countries. We then assess key characteristics of fee policies related to hospitals in selected countries for their contribution to equity and revenue. Next, we discuss important characteristics of risk-sharing programs in selected countries. After setting out some examples of the effect of alternative financing policies on hospital and sectoral efficiency and equity, we provide a policy summary of the suggested financing alternatives.

Revenue Performance

Several aspects of the effectiveness of alternative financing policies in generating revenues are considered. First, we assess the effect on revenue of such policies in different countries by examining available data on autonomous revenues (revenues raised through alternative financing mechanisms) as a share of total hospital revenue or expenditure. Second, we present available evidence on the potential of fees to recover costs by comparing charges with average costs. We also compare fee levels with per capita national income as an indication of the burden of hospital charges. Finally, we compare actual autonomous revenues with "potential" autonomous revenues, that is, the revenue that would have been collected if all those liable for reimbursing the hospital actually paid. Analysis of these revenue aspects of alternative financing programs helps to illuminate the relation between specific policies (such as fee retention, periodic adjustment of fees, and exemptions) and the extent to which a program generates revenue.

Cost-Recovery Ratios

The revenue effect of alternative financing programs in different hospitals may be compared by examining the cost-recovery ratio, defined as the ratio of total autonomous revenues (H_A) to total recurrent expenditures (C), H_A/C. A number of studies either calculate these ratios directly or provide data from which they can be calculated. The findings from these studies are summarized in table 5-1. The denominator of the ratios may be defined variously as actual or budgeted expenditures, depending on the available data. For our purpose, which is to get a rough sense of the impact of cost-recovery efforts on hospital financing, comparability of the ratios is not greatly affected by these alternative definitions. Additionally, there are other variations in the definition of the denominator of cost-recovery ratios that are more substantial. For example, the denominator in some cases is given as the total government subsidy rather than total hospital expenditures. In these cases, the figures presented in table 5-1 are based on a recalculation of the cost-recovery ratio, (if data are available to do so) to provide a consistent measure.

Several studies present partial cost-recovery ratios, which compare autonomous revenues to nonpersonnel recurrent costs. There are several reasons for this. One is that the explicit policy goal of some countries is to have hospitals recover these nonstaff recurrent costs through fees or other measures. In these countries it is often the case that personnel expenditures are made centrally and do not even appear in hospital budgets. Another reason is that, in some countries, hospitals cannot spend autonomous revenues on personnel. A third explanation is that, in any year, a hospital's staff may be a relatively "fixed" recurrent cost, whereas other recurrent cost items tend to be "variable" (that is, as use of services increases, the total cost of these items tends to increase as well). For those countries for which the appropriate data are available, table 5-1 also includes the cost-recovery ratio with nonstaff costs used as a base.

In most developing countries, public hospital cost recovery is very low and often zero. As can be seen in the table, hospitals in many countries that have user charges recover less than 15 percent of recurrent expenditures. However, there are examples of hospitals—such as most Chinese hospitals, two of Zaire's reference (district) hospitals, and Jordan's University Hospital—that derive most of their recurrent resources from these sources. In addition, cost-recovery ratios for hospitals within the same country may also vary widely, as in Mali, Bolivia, Jordan, and Turkey. Surprisingly, cost-recovery performance appears to be somewhat independent of a country's level of per capita income. For example, hospitals in some low-income countries (China and Zaire) and middle-income countries (Brazil and Korea; although not indicated in the table, these countries finance their hospitals entirely through charges and insurance programs) have very high ratios, whereas in other low- and middle-income countries, cost-recovery percentages are low (Lesotho, Honduras, Jamaica, Papua New Guinea) or moderate (Ethiopia, Indonesia, Mali, Jordan, Turkey). Distinguishing characteristics of each of the countries that have very high cost-recovery percentages are high charges in relation to costs and the presence of a fairly widespread insurance system designed to reimburse hospitals the costs of providing services to the insured.

Fees and Unit Costs

Another important indicator of the revenue performance of user charges for hospital services is the level of service fees in relation to the cost of service provision. In most developing countries, casual observation suggests that charges are well below average and even marginal costs. This has been difficult to document conclusively, however, because so few hospital cost studies have been performed.

Table 5-1. Cost Recovery Ratios in Selected Developing Countries
(percent)

Country and year of data	*Level of hospital*	*(H_A/C)*[a]	*(H_A/C)*[b] *(nonstaff)*	*Number of hospitals in study*	*Source*
Low-income countries					
China, 1986	I	90.1	118.9	8	Barnum 1989
	II	87.9	123.2	11	
	III	97.3	131.9	7	
Ethiopia, 1984–85	Urban	32.1	74.0	8	Donaldson and Dunlop 1987
	Rural	22.9	55.9	10	
Indonesia, 1985–86	All public	19.9		—	Prescott 1991
Mali, 1986	I	7.5	19.0	1	MPHSA 1987a, 1987b
	I		42.4	1	
	III		117.7	1	
Niger, 1986–87	I	14.8		1	Wong 1989
Zaire, 1988	III	78.9	168.7	1	Shepard, Vian, and Kleinau 1990
	III	66.3	100.5	1	
Middle-income countries					
Bolivia, 1988[c]	La Paz	47.5		6	
	Cochabamba	51.5		4	
	Santa Cruz	38.4		5	
Dominican Republic, 1986	I	2.7	7.9	1	Lewis 1987
	II	2.6	7.3	5	
	III	1.5	1.9	3	
Honduras, 1985	I	3.5	7.3	1	Overholt 1987
	II	4.5	10.6	6	
	III	5.3	14.4	8	

Jamaica, 1986–87	All public[d]	2.8	5.7	22	Kutzin 1989
	University	7.5		1	
Jordan, 1987	All MOH	13.3	38.5	—	
	All RMS e	22.3	68.7	—	
	University	51.3		1	
	All public[f]	22.6	52.5	—	
Papua New Guinea, 1985	All public	2.5		—	Thomason and others 1987
Senegal, 1985[g]	II		14.5	1	Hôpital de Saint-Louis 1986
St. Lucia, 1986–87	II	2.4	7.7	1	Russell, Gwynne, and Trisolini 1988
Swaziland, 1988–89	MOH	4.7		4	Collins 1990
	Mission	12.6		2	
Turkey, 1987	All MOH	12.6		—	
	All university	45.4		—	
Zimbabwe, 1989	I	7.3		4	Hecht, Overholt, and Holmberg 1992
	II	3.1		8	
	III	1.9		30	
Total Health Services[h]					
Botswana, 1983[i]	All MOH	1.3			
China, 1987	All public health	85.6			Burngarner 1992
Ghana, 1987[j]	All MOH	11.8			Waddington and Enyimayew 1989
Lesotho, 1986–87	All MOH	5.8			United Medical Enterprises 1988
Yemen Arab Republic, 1983	All MOH	3.3			

— Not available.

a. Total autonomous revenues (H_A) divided by total recurrent expenditures or revenues (C).

b. Total autonomous revenues divided by nonstaff recurrent expenditures or revenues.

c. Unweighted averages of the cost-recovery ratios of the hospitals in each region.

(Table continues on the following page.)

Table 5-1 (continued)

d. Fee collections as a percentage of total hospital expenditures, including drugs and medical supplies allocated to hospitals from central stores.

e. Royal Medical Society (these are also public sector hospitals).

f. The nonstaff ratio excludes University Hospital.

g. Official fees only; multiplied 6-month collection figure by 2 to match expenditure data.

h. Ratios reflect total hospital and nonhospital fee collections as a percentage of total public health expenditure. In most cases, the bulk of fee revenues come from hospitals, so these ratios understate hospital cost-recovery performance.

i. Uncertain if total expenditure includes retained fee revenues (if any).

j. Uncertain if the total MOH expenditure data include expenditures out of autonomous revenues.

Source: World Bank sector reviews and appraisal reports, except as noted.

There are several technical difficulties in making a comparison of fees and unit costs. First, in an examination of the appropriateness of the fee in relation to resources used, the figures being compared should be the fee and the marginal cost. With the exception of the few econometric studies summarized in chapter 3, however, marginal costs are unknown, and average costs must be used as an approximation. A second difficulty is that the definition of the service to which the fee applies may not coincide exactly with the definition used for cost estimates. For example, if the fee schedule includes prices for inpatient and outpatient care and separate fees for specific services such as drugs, X rays, or laboratory tests, the inpatient and outpatient prices should not be compared with the fully allocated average costs such as those given in table 3-1, which include all support services. Instead, fees should be compared with partially allocated average costs that exclude the costs of the intermediate services for which separate fees are charged. In table 5-2, which contains a summary of comparisons of fees and costs for a number of countries, average costs are adjusted to reflect the definitions of the services priced.

Conclusions drawn from table 5-2 are tempered by the small number of hospitals and countries for which data are available. For these few countries, it can be seen that fees for most services are well below average costs. In effect, governments have limited the financial liability of individuals for hospital costs and have, thus, implicitly provided protection (insurance) against the costs of illness through the budget subsidy. In none of these countries were fee levels explicitly related to the actual costs of providing a specific service as a matter of policy. Apparently, prices have been set based on political criteria, such as policymakers' perception of the population's willingness to pay. As a result, the fee-cost percentage for relatively low-cost services, such as laboratory tests, tends to be much greater than that for expensive services, such as inpatient ward care. It is also important to note that, with the exception of Niamey Hospital in Niger, although the fee for private inpatients is often much higher than for public patients in these hospitals, private patients also receive subsidized services.

Fees and Ability to Pay

The first pricing principle delineated in chapter 4 was that fees should be consistent with ability to pay and should not prevent essential access to health care. A rough indication of the burden of fees can be obtained by comparing fees and per capita GNP. A selection of public hospital user charges as a percentage of GNP per capita for various countries is given in table 5-3. Conclusions drawn from the table are only suggestive because the definition of services covered by the fees varies across

Table 5-2. Hospital User Charges or Reimbursement Rates as a Percentage of Average Recurrent Costs, Selected Countries

Country and year of data	Level of hospital	Inpatient: Public, per admission	Inpatient: Public, per day	Inpatient: Private, per day	Casualty/out-patient visit	Pharmacy prescription	Operating theater operation	Laboratory test	X-ray exam
Ethiopia, 1985[a]		78						19	
Grenada, 1985[b]	II	0	0	6	0	8	48	49	32
Honduras, 1985[c]	II	9			4				
	II	9			8				
	II	12							
	III	5			9				
Indonesia, 1987[d]	II		18						
	III		23						
	III-		16						
Jamaica, 1985–86[e]	I	2		46	16	23	13		24
	I	3		40	16	20	20	153	16
	II	5		60	27	28	33	165	31
	II	8		86	24	34	26		43
	III	4		53	20	45	17		142
Niger, 1986–87[f]	I	8		110	51			136	146
Papua New Guinea, 1988[g]	I	2		62	12			22	11
	II	2		24	26			23	9
	III	2		32	13			17	5
St. Lucia, 1986–87[h]	II		18	69	66		51	108	58
Zimbabwe, 1987[i]	I		44	48					
	II/III		36						

a. Figures reflect comparison of the median charge in fifteen Ethiopian hospitals to the estimated marginal cost of producing the indicated services. The authors also found that the median fee as a percentage of the marginal cost of a delivery was 9 percent.

b. There are no charges for public inpatients. The calculation of private inpatient costs includes intermediate services but not administrative overhead and surgical costs. However, since private inpatients are not charged for intermediate services (except surgery), the private inpatient fee/cost percentage is understated only by the lack of administrative costs in the average cost estimate.

c. Reported figures are simple averages of the percentage of average cost across four inpatient services in each hospital, with the greatest percentage occurring in obstetrics-gynecology and the lowest in medicine and pediatrics. According to Overholt, the average costs reported by Honduran hospitals are probably understated because they do not include ancillary services. The order of magnitude of the reported percentages is probably accurate, however.

d. Calculations are based on reimbursement levels from the public health insurance program (ASKES) for current and former civil servants and their dependents (spouse and up to three children). The ASKES reimbursement packet per diem is inclusive of nearly all inpatient services (only select special services, such as intensive care, heart operations, and hemodialysis are billed separately). The calculations use the median of inpatient packet reimbursement amounts for level II and III hospitals. In the table, level III- refers to Indonesia's class D hospitals.

e. Partially allocated inpatient and outpatient and fully allocated intermediate service average costs are used to be comparable with service fees. Average costs do not differentiate public from private patients.

f. Charge for public inpatients is inclusive of all services and is thus compared with average costs including diagnostic and therapeutic services. There are separate charges for specific services provided to private patients, so the bed fee is compared with the direct cost of these beds excluding specific services (weighted average of fees and costs for three different levels of amenity beds). Laboratory figure reflects public fee and the average cost for hematology test.

g. Fully allocated inpatient and outpatient costs are compared with inpatient and outpatient fees.

h. Partially allocated inpatient and outpatient and fully allocated intermediate service average costs are used to be comparable with service fees.

i. Fees for specific services (operations, injections, and so on) exist, whereas average costs include all components of inpatient stay. Thus, inpatient fee/average cost ratio is understated (percentage calculated using average fee of Z$5–35 range). Fees for children are lower than indicated. Private inpatient fee/cost ratio for central hospitals uses estimated average cost per day for private patients at Parirenyatwa Central Hospital of Z$125.

Source: Bitran-Dicowsky and Dunlop 1989 (Ethiopia); Mohr 1986 (Grenada); Overholt 1987 (Honduras); Prescott 1991 (Indonesia); Kutzin 1989 (Jamaica); Wong 1989 (Niger); JSI 1990 (Papua New Guinea); Russell, Gwynne, and Trisolini 1988 (St. Lucia); Hecht 1992 (Zimbabwe).

Table 5-3. Hospital User Charges or Reimbursement Rates as a Percentage of per Capita GNP, *Selected Countries and Public Hospitals*

Country and year of data	*Level of hospital*	*Inpatient*			*Casualty/ outpatatient visit*	*Pharmacy prescription*	*Operating theater operation*	*Laboratory test*	*X-ray exam*	*Maternity delivery*
		Public, per admission	*Public, per day*	*Private, per day*						
Ethiopia, 1985–86[a]	Urban		1.2	5.9	2.0		106.2	2.0	6.7	20.3
	Rural		0.8	2.1	1.0		21.2	1.1	10.4	5.4
Ghana, 1985[b]	II, III		0.4	0.9	0.2	at cost	2.0	0.1	0.6	0.4
Haiti, 1988[c]	University		0.8	29.2	0.2		13.9	0.3	0.8	1.4
	II, III		0.6	2.5	0.1		11.1	0.3	1.1	1.1
Honduras, 1986	I[d]	2.2			0.1			3.3	4.9	0.6
	II[e]	1.4			0.1			0.4	1.0	1.3
	III[e]	1.6			0.2			0.3	1.2	1.6
India, 1987	Public		0.03	1.7	free				0.3	
Jamaica, 1990[f]	Public	0.3		0.5	0.05	0.05	0.6	0.1	0.1	0.5
Mali, 1987[g]	I		0.5	2.8	0.7		13.9			
Nepal, 1986[h]	Public	free		3.8	0.1				1.1	
Niger, 1986–87[i]	I		2.3	2.2	1.3			1.3	4.8	
Papua New Guinea, 1988[j]	Public	0.6		1.7	0.1	free		0.1	0.1	free
Senegal, 1987	Official[k]		0.3	1.4	0.3					
	APH 1		0.1		0.1			0.2	0.5	1.1
St. Lucia, 1986–87	II		0.4	2.0	0.4	0.2	7.4	0.3	0.6	6.1
Swaziland, 1988[m]	MOH		0.1	1.0	0.1	0.1	0.5	0.1	0.1	0.1
Turkey, 1988[n]	MOH		0.1	2.8	0.1		12.0	0.1	0.2	
Zaire, 1988[o]	III	6.2								4.5
Zimbabwe, 1990[p]	I		1.0	4.1	0.7		3.7	0.8		1.7
	II		0.7	2.4	0.2		3.7	0.5	1.2	1.0
	III		0.7	2.4	0.1	0.3	3.7	0.5	1.2	0.3

a. Urban figures are based on the range of fees at several Addis Ababa hospitals; rural figures are based on the range of fees at several rural Ethiopian hospitals.

b. Inpatient and outpatient refer to adult fees; the inpatient fee used is the charge for an open ward with catering. Used the mean of a range of laboratory fees; the operation fee used is the mean of three fee levels.

c. Private bed per diem calculated as the mean of the endpoints of a range of fee levels. The operation fee is calculated as the mean of minor and major fees.

d. Outpatient department charge is the mean of walk-in and emergency fees. Laboratory fees are 1 lempira (L) for basic tests, and L5–150 for complex exams. X-ray charges are L5–150.

e. Based on the range of fees at several hospitals.

f. General ward fee is per admission; private patient fee is per day. Operation fee is the average of a range of five fee levels.

g. Fees are lower for civil servants. Private fee is the average of a range of amenity levels.

h. X-ray fee applies only to patients in paying beds.

i. Private bed fee is based on the weighted average of three categories of amenity beds. Laboratory figure reflects the public fee for a hematology test.

j. Higher specific service fees apply to private patients.

k. Adult fees only. Fees are lower for civil servants. Private fee is the average of a range of amenity levels.

l. Fees applicable to some of those exempt from official fees. Inpatient fee is CFA100/day for the first ten days, maximum of 500 for the days thereafter (the maximum total payment is 1,500).

m. Inpatient per diem fee based on the charge that applies to the first ten days in the hospital. Private inpatient fee is for a one or two bed room. Surgical theater fee is the mean of fees for short and long operations.

n. Laboratory test fee is for a blood cell count. Operation fee is the average of four fee levels; private per diem is the average of three amenity levels.

o. Inpatient fee covers all services; it is calculated as the mean of the high and low ends of a range of fees. Delivery fee percentage is calculated using the fee for maternity without prenatal care.

p. For level I hospitals, the public ward per diem and outpatient consultation fees reflect the midpoint of the income range of adult fees. For level II and III hospitals, the public ward fee is the per diem charge for the first fourteen days of an adult stay. For level I and II hospitals, the maternity fee is the average of per diem fees, and for level III, it is the average of the flat fee for booked and unbooked cases in the maternity ward. Operating theater fee is the average of major and minor operations, plus the fee for thirty minutes of general anesthesia for the middle-income range and excludes surgeon's fees. Laboratory fee is the charge to middle-income patients for a full blood count; X-ray fee is the charge to middle-income patients for a chest X-ray. Pharmacy charge is the inpatient daily charge for drugs in level III hospitals.

Source: Donaldson and Dunlop 1987 (Ethiopia); Government of Ghana 1985 (Ghana); World Bank data (Haiti); Overholt 1987 (Honduras); Seth and Gupta 1987 (India); Kutzin 1989 (Jamaica); Vogel 1988 (Mali; Senegal, official); Pande 1987 (Nepal); Wong 1989 (Niger); JSI 1990 (Papua New Guinea); APH 1987 (Senegal, APH); Russell, Gwynne, and Trisolini 1988 and Mohr 1986 (St. Lucia); Collins 1990 (Swaziland); World Bank data (Turkey); Shepard, Vian, and Kleinau 1990 (Zaire); GOZMOH 1991 (Zimbabwe).

countries. For example, in some hospitals inpatient fees cover all associated costs, such as pharmaceuticals and laboratory tests, whereas in others supporting services are priced separately. Also, fee policy is not necessarily uniform within countries, and in most cases the data given represent information from only a few hospitals. Nevertheless, the overall evidence indicates that fees are modest in scale in relation to per capita GNP in most countries. The median of fees is less than 1 percent GNPN per inpatient day and 0.2 percent GNPN per outpatient visit.

The scale of the median fee level can be placed in perspective by using an example of a family of five whose household income is 350 percent GNPN and who uses health services to a moderately high degree. If the family experiences one hospitalization of twelve days and five outpatient visits, total hospital expenditure will be 13 percent GNPN, or 13/350 = 4 percent of household income. Given that hospitalization is a relatively rare occurrence for a family, this is a modest burden on the average household.

This discussion only gives a suggestion that, in many countries in which there is some effort to recover cost in public hospitals, the scale of fees does not constitute a heavy burden in relation to income for the average household. This is not to suggest that the burden imposed on poor households by a sickness requiring hospitalization of a family member is negligible. Other costs, such as those associated with travel to the hospital, medical expenditures on behalf of the sick person prior to hospitalization, and the loss of the sick person's earnings or household labor may be quite significant. Hospital fees are thus one component of the overall financial burden of hospitalization. A full consideration of the burden of fees must consider the distribution of income and the effect of fee policy in public hospitals on the poorest groups in the population. The relation of fees to equity is considered further below in the section reviewing fee exemption policy.

Actual in Contrast to Potential Revenues

The revenues actually collected from hospital user charges are generally less than those that would have been collected if all those who used hospital services paid the designated amount. The principal reason for this is that some of those using the hospital are exempted from payment. Additionally, enforcement of payment is often lax, and nonexempt patients are able to receive hospital services without paying for them. Multiplying the total number of units of each service for a given year by their respective prices gives one measure of the potential revenue that could be obtained in the absence of any exemptions or evasion. Estimates of potential revenue obtained in this manner may be overstated, because having to pay the indicated fee might have dissuaded some persons from using the hospitals. This error would probably be small, however, given

the evidence of low-demand elasticities that exist at the low levels of prices found in many countries (see table 4-2). An alternative measure that gives some insight into the importance of lax payment enforcement is to calculate potential revenues only on the basis of the number of nonexempt units provided. Comparing actual with potential revenues allows an assessment of the performance of alternative financing mechanisms with given prices (table 5-4). That is, the effects of nonprice elements of alternative financing programs can be analyzed. The ratio of actual to potential revenues under the first definition can help illustrate the effect of formal exemption policy on total revenues, whereas that under the second definition can illuminate the effects of nonprice aspects other than exemption policy on the revenue performance of alternative financing policies.

Information on actual and potential revenues is available for only a small number of countries. Hospitals usually do not have records matching fee collection to the use of services, and far less frequently do they track the percentage of total patients who are officially exempt. The lack of such basic management information creates difficulties in assessing how well systems of user fees are operating. Of the countries cited in table 5-4, only hospitals in Honduras, Mali, and St. Lucia keep such records. Studies from these countries reveal that more than 20 percent of inpatients were exempted from charges at the Hospital Escuela in Honduras in 1985, nearly 70 percent of patient-days had to be provided free at the Hôpital du Point G in Mali in 1986 (Vogel 1988), and in St. Lucia, it was estimated that 93 percent of the population was statutorily exempt (Russell, Gwynne, and Trisolini 1988). Waddington and Enyimayew (1990) report that in Ghana in 1986, the revenue foregone because of exemptions (primarily for Ministry of Health staff and their dependents) was estimated to be equal to about 21 percent of total fee revenue for that year.

The limited data suggest that the wide application of official exemptions can considerably reduce total autonomous revenues, as indicated by the differences in the two ratios for the hospitals in Mali, Niger, and St. Lucia and the revenue lost to exemptions in Ghana. In each case, total revenues were substantially limited by rules granting a large number of persons exemptions from payment, including free services to those exempt for reasons other than their level of income. Hospitals in Honduras and Papua New Guinea demonstrate another finding, which may be relevant for many countries: many persons who are not officially exempt receive services without paying. The percentage of nonexempt nonpaying patients varies across countries and hospitals and is affected by many factors, which are described later in this chapter.

Available evidence on the revenue performance of alternative financing measures in hospitals in developing countries is limited but suggests considerable cross-country and within-country cross-hospital variation

Table 5-4. Actual as a Percentage of Potential Autonomous Hospital Revenues in Selected Countries

Country	*Level of hospital*	*Number of hospitals*	*Year of data*	*Fee retention rate*	*Percentage based on total revenues*	*Percentage based on nonexempt revenues*
Ethiopia[a]	Urban inpatient	3	1984–85	0	68	
	Urban outpatient	4	1984–85	0	90	
	Rural inpatient	6	1984–85	0	85	
	Rural outpatient	5	1984–85	0	73	
Honduras[b]	I (inpatient)	1	1985	100	22	28
Jamaica[c]	All MOH	22	1985	0	42	
	All MOH	22	1987	50	47	
Malaysia[d]	All public		1981			61
	All public		1982			62
	All public		1983			68
	All public		1984			74
	All public		1985			65
Mali[e]	I (inpatient)	1	1986	100	22	52
Niger[f]	I (inpatient)	1	1987	0	59	86
	I (outpatient)	1		50	40	68
St. Lucia[g]	II	1	1986–87	0	5	73
Papua New Guinea[h]	I	1	1988	0		48
	II	1	1988	0		35

	III	3	1988	0	16
Zimbabwe[i]	I	4	1989	0	62
	II	8	1989	0	57
	III	30	1989	0	79

a. Figures for Ethiopia reflect the percentage of patients who did not receive a "free care" card and thus paid for their care, which should give a good proxy for the percentage of potential revenue actually collected. Rural hospitals exhibited a wide range of total patients treated for free, from 3 to 60 percent of outpatients and from 9 to 41 percent of inpatients.

b. The fee per discharge at this hospital was either 35 lempiras or 10 plus 1 pint of blood. Potential revenues were calculated by multiplying the mean lempira fee (22.5) by the number of total discharges and the estimated (based on a three-month survey) number of nonexempt discharges.

c. An upper-bound estimate is that 30 percent of patients at Jamaica's public hospitals are exempt by statute. Estimates of actual/potential percentages for all MOH hospitals are overstated because statistics for the use of laboratory tests and drugs were not available to include in the estimate of potential revenue.

d. Malaysia's fee retention policy is unknown.

e. Figures reflect percentages for bed-day fees only.

f. Figures derived from a survey of hospital patients (594 outpatients, 379 inpatients). Figures in the nonexempt column reflect percentage of persons in the category, "Patients without Exemption from Payment," who paid for their care. Other categories of patients who were supposed to be fully or partially exempted from fees were found to have actually paid for their care.

g. St. Lucia's MOH has estimated that 92.7 percent of the population is legally exempt.

h. Potential revenue estimates adjusted for exempt services and services rendered to private patients.

i. Figures are a weighted mean of each hospital's fee collections as a percentage of its billings.

Source: Donaldson and Dunlop 1987 (Ethiopia); Overholt 1987 (Honduras); Kutzin 1989 (Jamaica); Ghazali and others 1987 (Malaysia); MPHSA 1987a (Mali); Weaver, Handou, and Mohamed 1990b (Niger); Russell, Gwynne, and Trisolini 1988 (St. Lucia); JSI 1990 (Papua New Guinea); Hecht, Overholt, and Holmberg 1992 (Zimbabwe).

in both cost recovery and the percentage of potential revenues actually collected. Hospitals in countries that recover more than half of their recurrent expenditures are typically associated with policies to limit government budget subsidies and implement large-scale insurance coverage. In most countries, cost-recovery ratios are much lower, largely as a result of low fees and nonpayment of fees.

Fee Policy in Selected Countries

The policy objectives concerning fees and the pricing principles described in chapter 4 provide an analytical context in which to examine the institutional characteristics of policies on user fees in a number of countries. The effect of alternative financing policies on revenue, equity, and efficiency depends on the level and structure of prices charged for specific hospital services. The available evidence suggests that revenues from fees should increase as hospital prices rise. This conclusion is based on studies (see table 4-2) that have demonstrated that, beginning from low levels of prices, the price elasticity of demand for public sector health services is low. Even in communities that have competing substitute services, the price elasticity, although somewhat higher, remains substantially less than one (Akin and others 1986; Heller 1982; Akin and others 1991). Achievement of the potential revenue, however, is strongly related to institutional arrangements for setting, collecting, and using fees. With regard to equity, most countries that have user charges for hospital care have either risk-sharing mechanisms or fee exemptions that attempt to shield most consumers from these prices. Several authors have raised concerns about the effect on equity of user fees, particularly if the fees are not graded to the income of the patient (Gertler, Locay, and Sanderson 1987; Gilson 1988; Waddington and Enyimayew, 1989 and 1990). Differences in pricing or formal or informal exemptions designed to protect the poor can be used to reduce the adverse equity consequences of financing programs with high user fees, but they must be carefully defined and enforced to be effective. Thus, demand elasticities and the income and other social characteristics of users set the environment that determines the effect of fees on revenue and equity. The success of meeting revenue objectives and safeguarding equity also depends strongly on the institutional arrangements for setting fees, adjusting fees through time, and the specific structure of fees. With regard to efficiency, most countries without large-scale insurance programs have relatively low fees. The incentives that very low fees provide to consumers and producers run counter to the achievement of sectoral efficiency, except for services characterized by significant market failures. The effect of alternative financing programs on the efficiency of resource use in the health sector is considered after the discussion of insurance later in this chapter.

Setting Fees and the Basis for Pricing

Fees may be set by central government, local governments, or individual hospitals. There is usually little information available on the marginal or even average costs of hospital services that could theoretically serve as a basis for pricing. For countries that use fees to supplement hospital operating budgets, fee levels are typically based on judgments (made either centrally or locally) of the population's ability to pay and an implicit assessment of the relative costs of hospital services. Countries that expect their hospitals to recover a significant percentage of recurrent costs (Brazil and Korea, for example) generally have insurance mechanisms to shield a large proportion of consumers from the full cost of hospital care, so the perceived ability to pay is less of an issue in determining hospital reimbursement rates. The ways in which these countries determine hospital fee levels vary.

There seems to be no discernible relation between a country's specific revenue objectives, such as modestly supplementing hospital budgets or substituting for a large portion of the government budget allocation, and the extent to which pricing decisions are decentralized. An exception to this may be that countries with comprehensive national health insurance programs tend to determine hospital reimbursement rates centrally. Table 5-5 relates the location of the pricing-reimbursement decision to the cost-recovery performance of a number of countries.

The role of either local or central government in the establishment of hospital prices ranges from determination of fee levels, to oversight of fees established by hospitals, to nearly complete laissez-faire. In situations in which pricing decisions are decentralized, the fees for similar services provided in different hospitals are likely to differ, even if the hospitals are of similar type. Centralized fee setting may or may not imply that one set of prices applies to all public hospitals in a country. Central policy may set fees that reflect regional differences in price levels; there may also be tolerance for autonomous deviations from central policy by individual institutions or local governments. Decentralized pricing is likely to generate fee levels that are more sensitive to the income levels and ability to pay in a specific hospital's catchment area. In addition, as is shown below, available evidence suggests that decentralized systems yield more timely adjustment of prices to keep pace with inflation. On the other hand, it can be politically difficult to justify differences in charges for the same services provided in similar public hospitals, particularly in a physically small country.

Differences in fees for similar services provided in hospitals of different types (for example, secondary or tertiary) may also prove difficult to justify, particularly if the referral system is weak. Although tertiary hospitals are intended to receive patients referred for advanced treatment, they often provide basic care to the population living nearby.

Table 5-5. Government Level That Sets Hospital Fees and Reimbursement Rates in Selected Countries

Hospital cost-recovery performance category	Centralized	Decentralized: Local government	Decentralized: Hospital
Low (less than 10 percent)	Botswana Ghana Grenada Jamaica Kenya Lesotho Papua New Guinea Senegal[a] St. Lucia Swaziland Zimbabwe		Dominican Republic Honduras Senegal[a]
Medium (10 to 30 percent)	Jordan Mali Turkey	Indonesia	Bolivia[b] Ethiopia
High (above 30 percent)	Brazil Korea	China[c]	China[c] Zaire

a. At the Hôpital Saint-Louis in Senegal, fee schedules set nationally and at the hospital are in use.

b. An official national hospital fee schedule exists in Bolivia, but in practice, hospitals set their own fees.

c. Chinese hospitals set their own fees, but these must be approved by local government price bureaus, which have established pricing guidelines.

Higher fees in tertiary hospitals for self-referred patients (bypass fees) may be justified only if there are viable lower-cost treatment settings available (see chapter 6 for a discussion of alternative treatment settings). Bypass fees could then be implemented to encourage use of these alternatives.

The information in table 5-5 indicates that revenue objectives may be addressed through use of either a centralized or decentralized method of price setting. For example, hospitals in Ghana, Jamaica, the Dominican Republic, and Honduras use some or all of the autonomous revenues to supplement their operating budgets, yet Ghana and Jamaica have a national fee schedule (Vogel 1988; Kutzin 1989), whereas hospitals in the latter two countries set their own fees (Lewis 1987; Overholt 1987). Similarly, in countries with somewhat higher hospital cost-recovery ratios, there are differences in how prices are set. Indonesia has national guidelines for hospital fees, but they are not binding on the levels of government that actually set prices (Prescott 1991). On the other hand, Mali's three national hospitals use a centrally determined fee schedule in an attempt to recover nonstaff operating costs (Vogel 1988). Turkey

also has an official, centrally determined fee schedule applicable to Ministry of Health hospitals, but the goal of these fees is to supplement hospital operating budgets. In addition, Turkey provides health services through a large social insurance program (covering about 42 percent of the population in 1987), which operates its own network of health facilities separate from those of the MOH. The MOH fee schedule does not apply in the hospitals of the social insurance program (World Bank 1990).

Countries with the highest hospital cost-recovery performance also show diversity in their methods of price setting. Pricing of hospital services in China and Zaire, as in the Dominican Republic and Honduras, is decentralized to the hospital level (Bumgarner 1992; Shepard, Vian, and Kleinau 1990), although prices set by Chinese hospitals are subject to the approval of local and provincial price bureaus. In Korea and Brazil, the bulk of hospital services is provided in private hospitals but financed by national health insurance programs. In these countries, the insurance program, as main purchaser of hospital services, determines the rates at which private hospitals will be reimbursed for treatment of insured persons. In Korea, hospital reimbursement rates for the insured population are determined in consultation with health service providers, and reimbursement and copayment rates are higher in tertiary care facilities than in community hospitals (World Bank 1989a). The central Brazilian National Institute for Medical Care and Social Security (INAMPS) sets payment rates for its case-based reimbursement system. According to INAMPS, the average values of bills for cases and procedures from all over the country are used as a basis for determining the appropriate payment rates (Rodrigues 1989b).

The decision to determine hospital fees centrally or in a decentralized manner depends on local conditions. For physically large countries or those with wide regional disparities in income levels, decentralized pricing may be more appropriate by being more responsive to local needs. In smaller countries or perhaps in countries with poorly developed referral facilities, it may be more appropriate for prices to be determined centrally. The relatively large number of countries in which fees are set centrally that recover less than 10 percent of expenditures through user fees suggests that decentralized pricing facilitates cost-recovery performance and is most appropriate when the revenue objective is to supplement hospital operating budgets. Centralized pricing often means that a political decision is needed to increase prices, and, as is shown in the next subsection, many countries have delayed fee adjustment for several years at a time.

Process for Adjusting Fees through Time

The regulations establishing a hospital fee schedule in many countries do not make formal provision for periodic adjustment of the fees as

needed to account for inflation. In countries in which periodic adjustment is not institutionalized, the decision to change fee levels must go through what is often a protracted political process. The experience of countries indicates that the regular adjustment of fee levels is probably the most important policy element contributing to the sustained ability of user fees to generate revenues through time.

The processes by which hospital fees are adjusted through time in several countries are summarized in table 5-6. In some countries, periodic adjustment of hospital fees is inherent in the fee structure, because fee levels are either directly related to cost (for example, Ghana's drug fee) or are tied to price or cost indexes (for example, the recommended fees in Indonesia). If fee setting is decentralized to the hospital level, as in China, the Dominican Republic, and Honduras, fees can be adjusted as needed by hospital managers to maintain or increase their real value. Thus, in countries in which fees are indexed or pricing is decentralized, fee adjustment is relatively easy and can be effected quickly to maintain real price levels. In countries in which there is a centrally set schedule of fees, adjustments occur through a political process and have proved difficult to implement. Examples of such countries are Jamaica, Lesotho, Mali, Papua New Guinea, Turkey, and Zimbabwe, each of which has experienced periods of several years during which nominal hospital fees were not changed. On the other hand, hospital reimbursement rates in Korea are defined through a political process, but this process occurs annually. Therefore, it is possible to institutionalize periodic fee adjustments through a centralized political process. Removing adjustments from such a process by tying fee levels explicitly to cost indexes, however, eases fee adjustment and probably improves the ability to maintain real fee levels.

Cost-recovery ratios are functions of real prices and real incomes. Therefore, policy with respect to periodic adjustment of fee levels is critical for maintaining autonomous revenues as a percentage of total hospital revenues or expenditures. One way to analyze the impact of this and other policies is to examine cost-recovery ratios for a given hospital or country for several years. Such longitudinal data on public health facilities in twelve countries are provided in table 5-7.

One caveat on the interpretation of these data is that because the cost-recovery ratios are indeed ratios, they are not solely a function of the level of autonomous revenues. The denominator of the ratios always involves, at least in part, the annual government subsidy to the hospital. This subsidy is determined through a political process, and its magnitude is constrained by budgetary limitations. Variation in the magnitude of the subsidy through time must be kept in mind when one interprets cost-recovery ratios. An increasing cost-recovery ratio may mean that public subsidies are being cut back. It does not necessarily imply that a hospital is doing a better job of raising revenues than it had been doing

Table 5-6. Means of Adjusting Hospital Fee Levels or Reimbursement Rates in Selected Developing Countries

Country	*Process by which fee levels or reimbursement rates are adjusted*
Low income	
China	Fees are set by hospitals, which can adjust them as needed over time, though local governments conduct some review and restrict the rate at which certain prices can increase.
Ghana	Fees are set by law and require a new act to change them. However, inherent in the current fee schedule is the means for drug fees to increase as the costs of these items increases; implementation of this feature has reportedly lagged.
Indonesia	Fees are set by the level of government that owns the hospital. However, the nonbinding fee guidelines established by central government tie fee levels to drug and food cost indexes, implying that fees adjust as costs change.
Lesotho	A fee schedule is set centrally for MOH facilities; it was not adjusted between 1980 and mid-1988. At that time, a new fee structure with higher charges became effective. It is unknown if the new fee schedule institutionalizes periodic adjustments.
Mali	No institutional fee adjustment mechanism exists. Current fees were set by a 1983 decree that defines the exact price for each kind of treatment. Another government act would be required to adjust these fee levels.
Zaire	Each health zone sets the level of prices for its reference hospital. Under the health insurance plan centered on the Bwamanda reference hospital, premiums and copayments have increased every year since 1986 to keep pace with inflation, as have prices charged to the uninsured.
Middle income	
Botswana	Government sets fee levels for MOH and mission facilities. Except for a modest increase in the private patient fee, prices were not adjusted in the period from 1972 to 1984. No data on fees since this time are available.
Dominican Republic	Hospitals set their own fees and are essentially free to adjust their prices over time; however, no fee may be charged for inpatient care.
Honduras	Hospitals set their own fees and are essentially free to adjust their prices over time; however, no fee is charged for pharmaceuticals.
Jamaica	Fees are set by an act of central government and require another such act to change them. Fees were increased in 1984 for the first time since 1968 and have not been raised since. There is no inherent means for prices to increase as costs increase.

(Table continues on the following page.)

Table 5-6 (continued)

Country	*Process by which fee levels or reimbursement rates are adjusted*
Jordan	Fees for MOH hospitals are set by central government, and government must act before prices can be adjusted. It is unknown how frequently government adjusts prices.
Republic of Korea	There is an institutionalized process for central government, in consultation with private providers, to periodically review and set reimbursement rates and copayment amounts.
Papua New Guinea	Public hospital fees are set by the central government, with no provision for increases to keep pace with inflation. Fee levels have not been increased since 1978.
Senegal	There is a two-tiered price system with official fees set by central government and retainable fees set locally by a private organization supporting the hospital (APH). Local fees can adjust as determined by the APH; it is unknown if national fees are subject to periodic review and adjustment.
Turkey	A central government act sets prices for MOH hospitals, and another such act is required to adjust them. Prices have been adjusted twice since mid-1981.
Uruguay	Central government sets the level of premiums paid to Mutual Aid Associations, which cover 40 percent of the population, including social security beneficiaries. The level of contributions is indexed.
Zimbabwe	Central government establishes a national fee schedule, and government action is required to adjust fee levels for most services. The fee schedule has not been adjusted since 1985. For diagnostic and treatment services, fees are based on a relative value scale (RVS) negotiated between the umbrella organization for private insurers and private providers. The RVS is regularly updated, and the public hospital fees for these services increase with each update.

Source: World Bank sources except as noted. Bumgarner 1992 (China); Vogel 1988 (Ghana, Mali, Senegal); Prescott 1991 (Indonesia); Shepard, Vian, and Kleinau 1990 (Zaire); Lewis 1987 (Dominican Republic); Overholt 1987 (Honduras); Kutzin 1989 (Jamaica); JSI 1990 (Papua New Guinea); Marquez 1989 (Uruguay); Hecht 1992 (Zimbabwe).

previously. For example, the budget allocation for Bolivia's Ministry of Social Welfare and Public Health fell by nearly 70 percent in real terms from 1980 to 1986 (World Bank data). This drastic decline was probably more responsible for the increase in the percentage of total hospital resources generated through cost-recovery mechanisms during this period than any changes in hospital service prices or improvement in fee collection practices.

This caveat notwithstanding, the table provides evidence of how important periodic adjustment of fees is to the amount of revenue

Table 5-7. Longitudinal Data on Cost-Recovery Ratios in Selected Countries
(percent)

Country	*Expenditure item*	*Year of data*	*Percentage cost recovery*	*Source*
Bolivia	Santa Cruz hospitals	1984	12.6	
		1986	64.0	
		1988	38.4	
Botswana	Recurrent MOH expenditures	1974	7.0	
		1978	2.9	
		1983	1.3	
China	All health expenditures	1985	77.8	Bumgarner 1992
		1986	81.9	
		1987	85.6	
Ghana	Recurrent MOH expenditures	1985	7.3	Waddington and Enyimayew 1989
		1986	7.8	
		1987	11.8	
Honduras	15 hospitals	1983	3.3	Overholt 1987
		1984	4.1	
		1985	4.2	
Indonesia	All government hospitals	1983–84	20.2	Prescott 1991
		1984–85	22.0	
		1985–86	19.9	
Jamaica	All hospital regions	1985–86	3.0	Kutzin 1989
		1986–87	2.8	
Lesotho[a]	Recurrent MOH expenditures	1974–75	16.0	
		1980–81	6.3	
		1986–87	5.8	
		1989–90	8.8	
Swaziland	All public hospitals	1983–84	2.0	Collins 1990
		1985–86	4.9	
		1988–89	4.7	
Turkey	All MOH hospitals	1984	9.1	
		1985	14.6	
		1986	14.4	
		1987	12.6	
Zaire	Bwamanda hospital	1985	48.7	Shepard, Vian, and Kleinau 1990
		1986	65.4	
		1988	78.9	
Zimbabwe	Recurrent MOH expenditures	1981–82	3.6	Hecht 1992
		1983–84	3.2	
		1984–85	4.3	
		1985–86	2.5	
		1987–88	3.0	

a. The 1989–90 percentage is an estimate based on actual revenues and budgeted expenditures for that year.

Source: World Bank sector reviews and appraisal reports, except as noted.

collected by alternative financing programs. Hospital fees in Botswana, Jamaica, Lesotho, and Turkey are not adjusted through an institutionalized process. In each case, fee levels have remained constant in nominal terms for years, whereas nominal hospital budget subsidies and expenditures have risen. In other words, prices for hospital services have fallen in real terms, causing a decline in the proportion of total revenues accounted for by fees. In Botswana and Lesotho, where fee levels went unchanged for ten years, this decline was dramatic. Lesotho's cost-recovery performance rebounded somewhat after an increase in fees during 1988. Jamaica is an interesting case because, in the latter year (1986–87), hospitals were entitled to retain 50 percent of the fees, a policy that was in place only during the last two months of the previous fiscal year. The change in retention policy provided an incentive to collect fees, but the drop in real prices (that is, the constant nominal fee levels) outweighed this effect and resulted in a slight decline in the percentage of the budget subsidy to hospitals collected as fees (Kutzin 1989).

Turkey provides another example of the importance of adjusting fees over time. Fee levels that had been in place since 1981 were increased at the beginning of 1985. The fee increase probably accounts for the increase in the cost-recovery ratio between 1984 and 1985. Fees were not adjusted during the next two years, however, and the cost-recovery ratio showed a decline (World Bank 1990). Similarly, in Swaziland, public hospital fees were increased in late 1984 for the first time since 1968, and the cost-recovery ratio for 1985–86 reflects the increase. Although fees were increased again in 1987, the magnitude of the increase was small, and the cost-recovery ratio has not risen above its 1985–86 level since this latest fee increase (Collins 1990).

Tying fee levels to price or cost indexes should be effective in maintaining real prices for hospital services, as nominal fees would tend to rise with inflation. Indonesia's nonbinding fee guidelines tie fees to indexes for drug costs and food costs, and cost-recovery data indicate that such a structure has been in place in most public hospitals. During a three-year period, the cost-recovery ratio remained about 20 percent, which suggests that real fee revenue stayed roughly constant (Prescott 1991).

Ghana implemented increased fees in mid-1985, which contributed to higher cost recovery in 1986 and 1987 (Vogel 1988). Revenues for the first six months of 1985 were based on lower fee levels. Fee collections during the latter months of 1985 were much greater than during the first half of the year, which suggests that the money-price elasticity of demand for services was quite low. The substantial improvement in 1987 may reflect improvement in the process of fee collection and an increase in patient volume. It is expected that the share of total revenues from drug fees would increase through time because these fees are tied to the costs of

drugs, whereas other fees are fixed in nominal terms (although increases in drug fees reportedly have been implemented only after a considerable lag). But on the whole, the percentage of cost recovery is expected to decline if fee levels for all services are not maintained at constant real levels, unless there is a substantial improvement in the percentage of potential revenue actually collected.

The available evidence suggests that hospital user charges generate revenue much more consistently in countries in which the level of charges can respond flexibly to economy-wide price changes. Such flexibility is most easily gained when prices are determined by individual hospitals, but it can also be built into a centrally determined national fee schedule by indexing fees or requiring that periodic adjustments be made. Evidence from several countries makes clear that it is often difficult to maintain constant real fee levels if an act of government is needed to change hospital charges. Introducing hospital user charges where none had existed before requires overcoming what are often considerable political obstacles. Designing a hospital fee system so that future price adjustments are removed from political processes is a key to maintaining the viability of a hospital fee system.

Structure of Hospital Fees

In addition to the level of charges in relation to costs, the extent of decentralization, and exemption and fee retention policies (discussed below), the structure of hospital fees, which involves the types of hospital fees and the units of service to which they apply, also affects the performance of a country's system of user charges. There are often incentives inherent in the structure of hospital fees or other reimbursement mechanisms that can have important effects on hospital efficiency (see the section on the effect of hospital financing programs on efficiency and equity, below, for a more detailed assessment of these incentives). Examination of the structure of fees within and across health facilities allows for insight into the extent to which a country's system of user charges provides proper signals for the use of health services and health sector resources (the second pricing principle described in chapter 4).

Fee structure varies across countries. Examples of structural variations in hospital fees include whether charges for inpatient services are per admission or per day; whether outpatient charges for first visits differ from charges for later visits; whether charges differ according to the type or level of hospital; whether diagnostic services are priced in detail; whether some defined services are excluded from fees; whether fee levels are related to the age or income of the patient; whether amenity services are present; and whether patients who refer themselves to a hospital for ambulatory care are charged more ("bypass" charges) than those who

are referred by a lower-level hospital or health center. Table 5-8 summarizes the structure of hospital fees in a number of developing countries.

Fees and facility type. Some countries (for example, Ghana, Korea, and Zimbabwe) set different fees according to the complexity of the facility. A penalty bypass fee for patients who refer themselves to hospital outpatient departments is also often recommended as a means to improve the economic efficiency of a country's public sector health facilities by encouraging patients to enter the health system at the least-cost provider. Pricing services in this manner does not prohibit self-referral, but a reduced government subsidy for this type of behavior would clearly signal the preferred referral pattern. Although there was no information on any of the countries studied that suggested the presence of such penalty fees for self-referred patients, the insurance plan for Zaire's Bwamanda health zone requires that all hospital cases must be referred from a health center (Shepard, Vian, and Kleinau 1990). A self-referred patient at the hospital is not entitled to the reduced rates charged to insured patients. This element of the program is clearly consistent with the pricing principle referred to above.

Price discrimination. The presence or absence of income-related price discrimination in the structure of fees is relevant to the principle that fees should be consistent with ability to pay. In most developing countries, equity concerns are addressed by fee exemptions for the poor rather than price discrimination built into the structure of fees (see discussion of exemptions and table 5-11). In countries with health insurance programs, the level of individual contributions is often based on a percentage of earnings. This implicitly provides for different levels of payment for health services according to income. Some countries, however, grade fees according to the income level of the patient. A formal sliding scale charge structure is most clearly demonstrated in Zimbabwe's four central hospitals, in which different inpatient per diem charges were, until recently, assigned to each of seven categories of patient income. According to a World Bank report (Hecht 1992), this attempt to build equity considerations into hospital pricing has not been very effective because the administrative skills required to apply such a detailed and rigid fee schedule successfully are often absent. As a result, the MOH has recently reduced the number of categories of patient income to three. An alternative approach is used in public hospitals in the Dominican Republic. Lewis (1987) reports that although hospitals set one level of fees, patients unable to pay the full amount often pay less, the ability to pay being ascertained by a social worker at the hospital. In effect, therefore, an informal sliding scale system is in place. On the basis of the information gathered from a small number of hospitals, Lewis concluded that, at

Table 5-8. The Structure of Hospital Fees in Selected Developing Countries

Country	Major aspects of hospital fee structure	Fees adjusted by: Patient age	Patient income	Amenity services	Hospital type
Low income					
China	Three factors affect the structure of fees: decentralized fee setting by hospitals; rigidities in the pricing system; and the stipulation that hospitals cover their nonstaff operating costs through fees. The result is that some prices approximate costs, some fall short of costs, and some exceed costs. In most hospitals, inpatient charges are per diem, and outpatient charges are higher for first than later visits. Most hospitals also have a detailed schedule of prices for specific diagnostic and therapeutic procedures.	No	No	No	No
Ghana	The fee schedule indicates that the fee charged for drugs should be the cost of the item. General consultation charges are lower in district hospitals than in regional and teaching hospitals. Charges for first visit specialist consultation in teaching hospitals are higher than for follow-up visits. Consultation charges in nonteaching hospitals are the same for each visit but are higher for adults than for children. There are somewhat detailed fees for specific diagnostic tests. There is a detailed list of surgical procedures grouped into three payment categories. Inpatient charges are per diem, with higher charges for adults than for children, two levels of amenity wards, and a low noncatering per diem option. About twenty-five (largely communicable) conditions are defined for which all fees except for drugs are waived.	Yes	No	Yes	Yes
Haiti	Charges for most services are higher at the University Hospital than at regional and district hospitals. Inpatient charges are on a per diem basis, with a range of private room fees much higher than fees for public wards.		No	Yes	Yes

(Table continues on the following page.)

Table 5-8 (continued)

Country	Major aspects of hospital fee structure	Fees adjusted by: Patient age	Patient income	Amenity services	Hospital type
Indonesia	The nonbinding guidelines for hospital tariffs define fees for outpatient visits and inpatient days that differ by level of hospital (for outpatient visits) and the amenity level selected (for inpatient days). Patients accommodated in wards with the least charges are also exempt from medical consultation and drug fees. Additional fees exist for surgery and diagnostic services. The insurance program for civil servants reimburses hospitals on a per diem basis a fixed price covering a bundle of services (room and board, supplies, diagnostic and therapeutic procedures). A few other items are billed separately. The insurance program for private employees also reimburses on a fixed per diem basis for a bundle of services, but at a much higher rate.			Yes	Yes
Kenya	Inpatient charges are per admission in MOH hospitals (for adults only); there are no outpatient charges. There are also charges for deliveries and some other services. A per diem charge exists for amenity beds. Charges are higher in the national hospital than in other MOH hospitals.		No	Yes	Yes
Mali	Inpatient charges are per diem and differentiated by amenity level. Patients with certain conditions (for example, cancer, psychiatric) are exempt from payment. There are also fees for outpatient consultations and diagnostic procedures.		Yes	Yes	
Niger	In all national hospitals and hospital centers, there are per diem fees for four levels of amenity beds and a per admission charge for public ward beds. Patients in amenity beds are subject to charges for diagnostic and treatment services, but this rule is not enforced. Ward bed patients pay a flat admission fee only. Amenity bed charges are lower for "public sector patients," a category that includes the military, civil servants, indigents, students, and employees of	Yes	Yes	Yes	No

	public enterprises. The general ward flat fee applies to inpatients who qualify on the basis of their household income. Children age five or under pay 25 percent of the adult fee; children age six to twelve pay 50 percent. The outpatient consultation fee has two levels: a lower fee for general practitioners, and a higher fee for specialists in specific services. Charges for medical acts and diagnostic examinations are highly differentiated and are based on a unit price for a category of services multiplied by a number assigned to the specific service. Unit prices are higher for private patients.		
Middle income			
Brazil	Hospitals (public and private) are reimbursed by INAMPS according to per procedure prospective prices, inclusive of all hospital services.	No	
Dominican Republic	Inpatient fees are prohibited, with a few minor exceptions. However, most hospitals either charge for blood when needed or require the patient to bring a donor. In addition, one hospital has a private wing that charges for all patient services and supplies. Outpatient fees vary widely across hospitals, with some charging for consultations and others not. All have some charges for outpatient laboratory and X-ray services, though the level of detail of these fees varies across hospitals.	Informal sliding scale	In one hospital
Honduras	There is some variation in fee levels across hospitals, but this is unrelated to complexity. Inpatient fees are generally per admission and include all services related to hospitalization except blood transfusions (though most hospitals accept blood as partial payment of the inpatient fee). Some hospitals, however, do charge on a per diem basis. There are fees for outpatient consultations as well as a range of charges for laboratory and X-ray procedures with varying degrees of differentiation according to resource intensity. There are no fees for drugs.		

(Table continues on the following page.)

Table 5-8 (continued)

Country	Major aspects of hospital fee structure	Fees adjusted by: Patient age	Patient income	Amenity services	Hospital type
Jamaica	For MOH hospitals, inpatient fees are per admission, with separate fees for services received during the inpatient stay. Private inpatient fees are per diem. There is a flat fee for deliveries and a five-tiered schedule of surgery fees related to complexity. Charges for laboratory and X-ray tests do not differentiate by the type of test. The University Hospital has a four-tiered structure of fees, with payment amounts defined by the income category of the patient. A very detailed price list for specific laboratory and X-ray procedures is also defined at this hospital.	No	Only in University Hospital	Private wards in a few hospitals	No
Korea	Hospitals are reimbursed on a fee-for-service basis according to the nationally derived reimbursement rates, which vary somewhat according to the type of hospital. Hospitals are paid for each itemized service. Copayment rates also vary according to the level of hospital, and, in addition, according to the level of patient indigence.	No	For the indigent	No	No
Papua New Guinea	Public inpatients are charged per admission. Many hospitals have "intermediate" (that is, amenity) wards, which charge per day. There are also higher charges for intermediate patients for outpatient, diagnostic, and treatment services.	No	No	Yes	No
Senegal	The national fee schedule defines inpatient per diem fees for adults and two age-differentiated groups of children across three amenity levels. There is also a charge defined for indigents. Inpatient fees also differ by civil service or military status of the patient. Outpatient visit fees differ by civil service status. Apparently, inpatient and outpatient fees are inclusive of all services provided by the hospital. The local (APH) fee schedule at the Hôpital Saint-Louis also has	Yes	Yes	Yes	

	a per diem inpatient charge, and the rate charged decreases by half after the first ten days. There is a consultation fee, a general laboratory fee, and four X-ray fees based on the type of test.				
Turkey	Inpatient fees in MOH hospitals are on a per diem basis, with four amenity levels. There are also four surgery fees, and there are a number of specific fees associated with outpatient consultations. There are also detailed lists of fees for X-ray, laboratory, and pharmaceutical services.		No	Yes	
Zimbabwe	For inpatient care at central hospitals, there is a per diem fee progressively related to three patient income (and insurance) categories (with fees for children about 60 percent of adult fees at each defined income level) for general wards. Inpatient charges are higher at provincial than at district hospitals and are not income-related in either, and the per diem rate is halved after fourteen days (children's ward fees are one-third of adult fees). There is a higher per diem rate for private ward care at all hospitals. At small rural hospitals, there is an inpatient flat fee that results in considerably lower charges than at higher levels. Consultation fees are differentiated by income and insurance status at central hospitals, but not elsewhere, and charges are graduated by level of hospital. There are fees for specific hospital treatments (for example, operations, anesthesia, plaster casts, occupational therapy) that are income-related but are not graduated by hospital level. Drug charges are income-related at the central and provincial levels and are based on private sector fees; at district hospitals, there is a flat fee. Laboratory charges are 1.4 times higher in central hospitals than at lower levels, and charges for these services are income-related in all hospitals. Radiology charges are income-related at all levels and are the same in provincial and district hospitals.	Yes	Yes	Yes	Yes

Source: World Bank sources, except as noted. Bumgarner 1992 (China); Government of Ghana 1985 (Ghana); Prescott 1991 (Indonesia); Vogel 1988 (Mali, Senegal); Weaver, Handou, and Mohamed 1990a (Niger); McGreevey 1988 (Brazil); Lewis 1987 (Dominican Republic); Overholt 1987 (Honduras); Kutzin 1989 (Jamaica); JSI 1990 (Papua New Guinea); GOZMOH 1991 (Zimbabwe).

most, hospitals fully exempt up to 60 percent of their patients and charge less than the full amount for up to about 50 percent. The application of an informal sliding scale at the facility level appears to be easier to manage than a formal system using specific charges for relatively narrow income categories.

Most of the countries studied offer hospital patients the option of paying a higher fee in exchange for inpatient services with nontreatment amenities, such as a semiprivate or private room. Ghanaian hospitals have four levels of charges for inpatients: two levels of amenity beds, a general ward charge, and a lower general ward charge for those not wishing catering services (Government of Ghana 1985). Other countries (for example, Haiti, Indonesia, Turkey) also define several amenity levels in their fee schedules. Providing amenity services is a way that hospitals can generate higher revenues from those willing to pay for them, and such revenues can possibly be used to cross-subsidize care provided to poorer patients. In many countries (for example, Grenada, Jamaica, Papua New Guinea, St. Lucia, and Zimbabwe), however, the net revenue arising from provision of these services is negative (see table 5-2). Thus, amenity services are often subsidized, and it is not always true that the margin (that is, revenues minus costs) for these services is much greater (less negative) than for general ward services. The limited evidence suggests that amenity services are not providing a source of profit to public hospitals and thus do not yield cross-subsidization possibilities. It is possible that they could do so, but the price charged to (and the amount collected from) private patients in public hospitals, especially for those covered by formal insurance programs, must first be raised.

Several other characteristics are commonly found in hospital fee schedules. For example, although hospitals in most countries do not build equity concerns into their fee structure by scaling charges to patient income, many charge lower fees for children than for adults. Another interesting characteristic of most of the fee systems studied is that they are very detailed, even in countries in which the level of fees is quite low. For example, Jamaica's fee schedule defines five different payment rates for surgical operations, Ghana has three surgery rates and specific charges for different laboratory and radiological examinations, and Turkey has four surgical rates and different charges for specific diagnostic investigations. If a country intends to keep fees at a very low level, the fee structure should be simplified to the degree possible in order to minimize administrative costs. Griffin (1987), using examples from Senegal and Sudan, argues that enforcement of only a single fee at admission is a feasible method for increasing revenues substantially in most hospitals in which cost-recovery objectives are modest.

Revenue performance of specific service fees. The contribution of fees for specific hospital services to total autonomous revenues depends princi-

pally on the level and structure of fees. Revenues are also dependent on the (nonpolicy) demand characteristics of each service (that is, underlying epidemiology and the income and price elasticities). Table 5-9 provides data from hospitals in five countries on the contribution of specific service fees to total autonomous revenues. Although the evidence is limited, some conclusions can be drawn.

China, which has achieved the greatest degree of cost recovery of the countries in the table, demonstrates the importance of fee structure and the level of fees. Although fee setting is decentralized to the hospital, rigidities in the pricing of certain services remain. The effect of these rigidities is that prices for services that were in use during the 1950s, such as bed fees and outpatient registration charges, are set below cost, whereas charges for newer, technologically advanced services, including many diagnostic and therapeutic procedures, can be set much higher in relation to cost (there are no price guidelines for these newer services). The fee for drugs is typically set about 10 to 20 percent above the cost to the hospital (Bumgarner 1992). The result of this fee structure is clearly shown in table 5-9. Fees for drugs account for more than 50 percent of total autonomous revenues, and registration and bed fees together account for less than 10 percent.

Drug fees also play a leading role in Ghanaian hospitals. The fee schedule indicates that the price for drugs should be equal to their cost. Even if this policy has not been fully implemented, it is likely that drugs are priced closer to their cost of provision than are other hospital services. Data reported by Vogel (1988) indicate that drug fees accounted for between 35 and 73 percent of total fee revenues in seven hospitals. Given the characteristics of the drug fee and the apparently greater willingness to pay for drugs than for other services, it is likely that drugs will account for an increasing share of total hospital fee revenues.

Prescription drugs may be characterized as a low-cost, high-volume hospital service. The evidence from China and Ghana suggests that charging at or above cost for this type of service can generate a substantial share of hospital revenues. The low cost of drugs means that they are affordable to many people, and the fact that people receive a tangible item seems to be associated with willingness to pay. A negative aspect of this situation that has manifested itself in China is that hospitals are overproviding drugs to patients. Bumgarner (1992) reports an average of 2.3 drugs prescribed per patient-visit in China. This overprovision of drugs not only leads to cost escalation but also appears to run counter to sound medical practice. In this example, the price signals to providers, contrary to the second pricing principle in chapter 4, are not generating incentives for proper use of resources.

Jamaican hospitals, which have a low level of cost recovery compared with that in hospitals in many other developing countries, show more balance in the contribution of different services to overall fee revenues.

Table 5-9. Share of Specific Services in Hospital Autonomous Revenues in Selected Developing Countries
(percent)

Country	*Level of hospital*	*Number of hospitals*	*Year of data*	*Bed fees*	*Registration, consultation*	*Drugs*	*Laboratory*	*X ray*	*Deliveries*	*Other*
China[a]	I	8	1986	6.3	1.0	51.2				41.5
	II	10	1986	6.3	1.7	57.6				34.5
	III	8	1986	4.8	1.3	64.8				29.1
Ghana	I	1[b]	1986	23.9	15.6	43.9	3.3	2.9		10.5
	I	1[c]	1986	19.2	27.8	35.1	4.0	1.4	0.6	11.8
	Accra Region	3[d]	1986	22.4	13.4	50.0	1.4	2.7	1.3	8.9
	Ashanti Region	2[e]	1986			55.9				
Honduras[f]	I	1[g]	1984–85	22.0	27.0		2.0	10.0		39.0
	II	3[h]	1982–85	50.0	20.7		12.0	14.3		2.7
	III	2[i]	1983–85	44.0	25.0		9.0	11.5		10.5
Jamaica[j]	I	1	1985–86	23.2	19.1	14.9	8.1	12.2	14.9	7.6
	II	4	1985–86	28.2	10.2	17.2	7.3	8.8	11.7	16.9
	III	14	1985–86	24.0	10.2	26.1	4.9	7.3	18.6	8.9
Mali	I	1[k]	1986	48.2	10.8		3.2	19.3	1.8	16.7
	I	1[l]	1986	23.6	43.7		0.5	6.1		26.0
	III	1[m]	1986	25.1	47.9		1.7	24.1		1.1
	National	3	1986	33.8	31.4		1.8	15.0		17.3

Niger[n]	I	1	1988	10.0	22.9	22.2	17.8	27.8
	II	4	1988	35.5	21.3	16.2	5.8	21.1
	III	3	1988	40.2	8.1	7.2	1.3	43.2

a. Specific fee data are available only for bed-days, outpatient registration, and drugs. All other specific service fees are grouped under "Other."

b. Based on two months of data. "Other" includes fees for operating theater, mortuary, dental, and physiotherapy services.

c. Based on one month of data. "Other" includes fees for ambulance, mortuary, and dental services.

d. Based on one month of data. "Other" includes fees for operating theater, dental, mortuary, and other services.

e. Based on one month of data. Data are available for drug and total fee revenues only.

f. "Other" also includes fees for the blood bank, electroencephalogram (EEG), electrocardiogram (EKG), special drugs, dental clinic, and health cards. Because blood donations can be made in lieu of part of the inpatient fee, Overholt lists the fees for the blood bank as an inpatient service.

g. The blood bank accounted for 30 percent of total revenues at this hospital.

h. Unweighted average of percentages of three hospitals. Inpatient percentage includes the blood bank, which could not be separated out in the data.

i. Unweighted average of percentages of two hospitals.

j. "Other" includes fees for morgue, dental, physiotherapy, and operating theater services.

k. "Other" includes fees for special examinations, nuclear medicine, drivers' licenses, EKG, and other services.

l. "Other" includes arrears and fees for private rooms and EKG.

m. "Other" includes fees for dental consultations.

n. "Other" includes revenues from employer contracts and outpatient fees for medical certificates, annual checkups, and other ancillary services. Bed fee reflects all revenues from inpatients; no disaggregation is available.

Source: Barnum 1989 (China); Vogel 1988 (Ghana); Overholt 1987 (Honduras); Lewis 1988 (Jamaica); MPHSA 1987a (Mali); Weaver, Handou, and Mohamed 1990a (Niger).

The relatively high percentages of revenues attributed to admission and delivery fees may be a function of the process for collecting these fees and, as a result, the extent to which fees for these services are collected in relation to other services. The fee collection process for inpatients generally requires a deposit, so it is harder for those who are not officially exempted to avoid payment than it is for nonexempt outpatients. Reflecting this are data from one tertiary hospital in 1986 (reported in Kutzin 1989), which indicate that the percentage of potential revenue (assuming no exemptions) actually collected was greater for maternity (59 percent) and admission (52 percent) fees than for other services.

The ability to derive substantial revenues from more costly and less frequent events (such as inpatient admissions, CAT scans, or surgical operations) is constrained by the extent of health insurance coverage. In China, insurance coverage of 20 to 35 percent of the population makes high charges feasible for some expensive procedures. None of the other countries in table 5-9 charges as near to cost as does China, nor do any of them have insurance programs that are as widespread. The limited evidence available suggests that countries that seek to supplement hospital revenues through fees but that are not interested in comprehensive explicit national health insurance programs could increase revenues by charging at or near average cost for services that are highly used but low in cost and that provide patients with a tangible output, such as drugs, laboratory tests, and routine X rays. Patients tend to perceive receipt of such services as an indicator of good quality, and their willingness to pay is consistent with the third pricing principle, which links fees and quality. However, prices set in this manner can generate signals to providers and consumers leading to improper service use. Therefore, monitoring of use is needed to check the incentive to overproduce profitable services. Although it is still appropriate to charge for expensive procedures and inpatient stays, these services should be subsidized for the uninsured. The suggested policy is, in effect, an implicit form of catastrophic health insurance, in which government covers some or most of the services defined as having catastrophic costs. More complicated systems would involve limiting, but not eliminating, individual or family financial liability for chronic conditions, the costs for which could become substantial in time.

Fee Retention Policy and the Efficiency of Revenue Collection

In chapter 4 it was hypothesized that the efficiency of revenue collection is linked to whether or not the hospital can retain a portion of the user fees it collects. Providing hospital managers a financial incentive to enforce fee collection is likely to be more effective and less costly than a bureaucratic solution that requires periodic oversight to certify that all

potential revenues are being collected. Furthermore, in countries in which most resources allocated to the hospital from government budgets are tied to narrowly defined expenditure categories, the use of fee revenues by hospital administrators gives local managers greater autonomy and responsibility for resource allocation decisions. In table 5-10 we describe hospital fee retention policy in a number of countries and, for those countries that allow for at least partial retention, the ways in which these autonomous revenues have been used.

For hospitals in countries that do not allow fee retention, fee revenues essentially substitute for government financing by replenishing the treasury with a portion of the revenue that had previously been allocated to the hospital. Such a system provides little incentive for hospital managers to make strong efforts to collect fees because the hospital receives no direct benefit from the revenues and essentially serves the role of tax collector. Local managers do not gain greater responsibility and autonomy under these conditions. In addition, allowing a degree of fee retention makes it possible to reduce the net government subsidy to hospitals if this policy induces an increase in actual revenues collected. The effectiveness of hospital fees as a cost-containment mechanism to ration demand, however, does not depend on whether revenues are retained, unless the lack of retention results in such little effort to collect that potential patients do not consider the official fees relevant.

Indirect evidence of the importance of fee retention is provided by joint consideration of the policies in table 5-10 and the cost-recovery ratios in table 5-1. Cost recovery is more sensitive to the level of prices than to retention policy, but an examination of this evidence is suggestive. In general, countries that disallow hospital fee retention tend to have low cost-recovery ratios, and countries with high cost-recovery ratios tend to allow full or partial fee retention. There are exceptions, however. The most striking are Ethiopian hospitals, which are not entitled to retain any fee revenues yet show moderately high cost-recovery ratios. It is also important to note that several countries that allow fee retention, such as the Dominican Republic, Honduras, and Jamaica, have low cost-recovery ratios. The expectation that achieving significant cost recovery requires that hospitals be allowed to retain some portion of fee revenues is given only weak support by the indirect evidence of cost-recovery ratios.

The best way to examine the effect of fee retention policy on the level of autonomous revenues is to compare the ratio of actual to potential revenues across countries with different retention rules or for one country for the years before and after a change in these rules. Unfortunately, the data are so limited that any conclusions drawn are merely suggestive. Indeed, for cross-country comparisons, table 5-4 shows that hospitals in Ethiopia, which must turn over all fee revenues to the treasury, collected

Table 5-10. Fee Retention and Use in Selected Developing Countries

Country	*Fee retention policy*	*Use of fee revenues*
Low income		
Burundi	All revenues are sent to the central treasury.	
China	Hospitals retain all fee revenues.	Intended to cover nonstaff operating costs. Surplus revenues are also used for medical equipment purchases and staff bonuses.
Ethiopia	All revenues are sent to the central treasury.	
Ghana	Hospitals retain 25%; the MOH is sent 25%, and the Ministry of Finance is sent 50%.	Unknown.
Indonesia	Revenue accrues to the level of government that owns the hospital and is sometimes returned to the hospital. Hospitals retain all revenues collected from the insurance scheme for civil servants.	Revenues returned to the hospital are in the form of budget allocations. No data are available on how retained revenues from insurance are spent.
Kenya	Beginning in late 1989, hospitals were entitled to retain 75% of collections; the remaining 25% is deposited with District Health Management Boards. The cost-recovery program was abandoned in late 1990.	Hospitals used their share for operating expenses; the district share was for reallocation to PHC and improvement of care in poorer areas. However, district boards lacked authority to perform this intended role.
Mali	Hospitals retain all fee revenues.	Uses vary among the three national hospitals, but principal among them are drugs, maintenance and repairs, and a category defined as "operating costs."

Country	*Fee retention policy*	*Use of fee revenues*
Niger	50% of revenue from office visits, medical acts, and diagnostic tests is retained by the hospital; the remainder is sent to the central treasury.	Retained revenues are distributed among hospital staff (the ristourne).
Middle income		
Brazil	Hospitals (which are mostly private) retain all of the reimbursement paid by INAMPS.	Hospitals use fees to cover all costs.
Dominican Republic	Hospitals retain all fee revenues, with virtually no restrictions on their use.	Drugs are the most frequently purchased item, followed by maintenance and medical supplies; several hospitals also allocate funds to hire unskilled staff or top off professional salaries.
Honduras	Hospitals remit fee revenues to the treasury but have the exclusive right to use the funds. Unexpended revenues revert to the treasury at the end of the fiscal year. The MOH must approve all expenditures.	There is considerable cross-hospital variation. Major categories of expenditure are general materials and supplies, casual labor and overtime, and surgical supplies.
Jamaica	All revenues are sent to the MOH, and 50% are returned to collecting hospital.	Most expenditure is on maintenance, followed by equipment and supplies.
Jordan	All revenues are retained by the National Medical Institution; it is uncertain if they are reallocated across hospitals.	Unknown.
Korea	Hospitals are private and retain all reimbursements and copayments.	Hospitals use fees to cover all costs.
Papua New Guinea	All revenues collected by hospitals go to general revenues.	

(Table continues on the following page.)

Table 5-10 (continued)

Country	*Fee retention policy*	*Use of fee revenues*
Senegal	Revenues from official fees are sent to the central treasury; revenues from local fees supplement the hospital budget.	Autonomous revenues from the first month under the local fees were spent predominantly on medications, maintenance, and day-to-day functions.
Swaziland	All revenues from MOH hospitals are sent to the central treasury; mission hospitals retain all revenue.	Mission hospitals use revenue for their operating expenses.
Turkey	MOH hospitals retain 85% of fee revenues, 5% is sent to the central treasury, and 10% is transferred to the MOH.	MOH hospitals can use their funds only to purchase equipment or cover operation and maintenance expenses. Data on two such hospitals showed the majority being devoted to drugs and medical supplies.
Zimbabwe	All fee revenues are sent to the central treasury.	

Source: World bank sources, except as noted. Bumgarner 1992 (China); Donaldson and Dunlop 1987 (Ethiopia); Vogel 1988 (Ghana, Mali, Senegal); Prescott 1991 (Indonesia); Epstein and Coultas 1991 (Kenya); Weaver, Handou, and Mohamed 1990b (Niger); McGreevey 1988 (Brazil); Lewis 1987 (Dominican Republic); Overholt 1987 (Honduras); Lewis and Parker 1991 (Jamaica); JSI 1990 (Papua New Guinea); Collins 1990 (Swaziland); Hecht 1992 (Zimbabwe).

the highest percentage of fees from their patients, much higher than hospitals in Honduras or Mali, which retain 100 percent of revenues. This outcome is contrary to expectations, and it may be due to the very long length of time that hospital fees have been in place in Ethiopia.

It is interesting to compare Jamaican hospitals in 1985, when they retained no fees, with the same hospitals in later years, when they retained 50 percent of the fees. Assuming that the proportion of patients exempt by statute was the same in each year, we find from table 5-4 that, for all MOH hospitals, the revenue collected per patient, at the same level of prices, was slightly greater in 1987 (with 50 percent retention) than in 1985 (with no retention). The data suggest that the allowance of fee retention increased collections but only to a small degree. Some support for the significance of fee retention is provided by Lewis (forthcoming), who suggests that the large increases in cost-recovery ratios between 1983–84 and 1985–86 and in revenues per patient between 1984–85 and

1987–88 in Jamaican hospitals demonstrate the effect of the incentive to retain 50 percent of the fees. This result, however, is obscured by the fact that some proportion of the increase in revenues was also due to the substantial increase in the level of fees in late 1984 for the first time since 1968.

Niger's Niamey National Hospital provides an interesting counterexample to the Jamaican case, showing a large effect on total revenues of specific fee retention provisions. Beginning in 1987, the "ristourne"—a system in which 50 percent of revenues from office visits and diagnostic examinations was distributed among hospital staff—was introduced. The hospital's 1988 activity report noted that revenues from these services increased by nearly 70 percent between 1986 and 1988 and that the share of total hospital fee revenues accounted for by these services increased from 45 to 74 percent (Weaver personal correspondence 1990). Prices remained constant during this period.

Evidence from public and mission hospitals in Swaziland (Collins 1990) supports the argument that fee retention policy yields increased efforts to collect fees. Since late 1984, public and mission hospitals have charged the same fees for public patients under a uniform hospital fee structure set by the government. Although they both charge the same fees, mission hospitals retain 100 percent of collected revenues, whereas public hospitals must return such revenues to the central treasury. Indirect evidence of the effect of the different fee retention policies is provided in table 5-1, in which cost-recovery ratios in two mission hospitals are shown to be substantially higher than in four public hospitals. More direct evidence is provided by comparing the unit revenues (that is, revenue per patient, in this case measured as the total revenue divided by the sum of inpatient days and first outpatient visits) in public and mission facilities. In 1988–89, revenue per patient was more than three times as great in a large mission hospital as in a large public hospital. Given that prices for services provided to public patients in both hospitals were the same, Collins (1990) suggests that higher unit revenues in mission hospitals could be explained by some combination of the following: higher fees for private patients, a greater number of ancillary services charged to outside private physicians, and better collection practices. Another possible reason for the difference would be that patients who used mission facilities had higher incomes than those who used public facilities. Collins concluded that retention policy has been an important factor contributing to better collection practices in mission hospitals.

Although the theoretical argument that individuals and institutions make a greater effort to collect fees if they benefit from the collection is strong, there is insufficient current empirical evidence to support this conclusion and to measure the effect of fee retention policy on total

autonomous revenue collections. If an estimate is to be made of the optimal fee retention rate, data on the ratio of actual to potential revenues with alternative retention policies must be gathered (see chapter 4, note 3). In addition, more evidence is needed on other factors that may also affect the efficiency of hospital fee collection activities, such as billing and banking practices. Despite this lack of strong direct empirical support, most countries with medium to high rates of cost recovery allow substantial if not complete retention of revenues by hospitals. The logic of the argument for allowing fee retention coupled with this indirect evidence suggests that collection incentives should be part of a system of hospital user charges.

Use of Fee Revenues

The third pricing principle delineated in chapter 4 was that fees and the quality of services should be linked. More specifically, given that establishing or raising the price of a service will reduce the quantity demanded, a portion of revenue should be retained by the collecting facility and used to improve the quality of services. This issue is particularly important when the revenue objective is to supplement the operating budget of public hospitals because such an objective implies an intention to improve the quality of hospital output by increasing the quantity of inputs. Policy simulations (described in chapter 6) from Kenya and Nigeria illustrate that improved quality (suggested by supply-side indicators, such as increased availability of drugs) purchased by these supplemental revenues can boost demand to a point that will more than offset the effect of the price increase. Unfortunately, little actual data are available on how fee revenues are used, and almost no data are available on the response of demand to changes in both price and quality.

The right-hand column of table 5-10 is a description of how fee revenues are dispensed in certain countries. Fees are intended to be used to supplement government budget allocations in the Dominican Republic, Honduras, Jamaica, Kenya, and, partially, in Senegal. Although the amount collected in relation to hospital expenditures in these countries is small, fee revenues can be important in maintaining services when other funds are insufficient. In Jamaica, for example, these revenues allowed hospital managers to make emergency maintenance expenditures that in previous years (prior to fee retention) would not have been made. In the past, insufficient funds to repair buildings and equipment have sometimes resulted in entire wards or other hospital units being shut down (Lewis and Parker 1991). Therefore, even though Jamaican hospitals recover an average of less than 6 percent of nonstaff operating costs through fees, these revenues have proved important at the margin.

Kenya's experience highlights the danger of not maintaining government budget allocations after hospitals have been entitled to retain fee revenues. Reforms introduced in December 1989 allowed hospitals to retain fee revenues (75 percent) for the first time; the remainder was given to district health management boards for reallocation to primary and preventive care in poor areas. The goal, perhaps based on the conclusions of the aforementioned policy simulations (Mwabu and Mwangi 1986), was to use fee revenues to supplement government budget allocations and thus improve the quality of hospital services. In fact, however, given diminishing availability of resources from the central government, the fee revenues served as a substitute for, rather than a supplement to, the budget allocation and reportedly added nothing to the quality and availability of services. Kenya's user fee program also suffered from poorly defined administrative procedures for patient exemptions and expenditure of revenues by the district health management boards. The combination of increased fees and no visible improvement in service quality, a poorly administered system, and growing political and economic difficulties led the government to abandon cost recovery in the health sector in September 1990 (Epstein and Coultas 1991). Thus, contrary to the third pricing principle in chapter 4, fees were raised but quality was not improved. This contributed to the disintegration of Kenya's user fee system.

Exemption Policy

An argument often made against the establishment of user fees for public hospital services is that fees will reduce access to needed services for low-income groups. Exemption from payment on the basis of income is a form of price discrimination and is consistent with the principle that fees should not prevent essential access to health care. Most countries that charge for the use of public hospitals explicitly define categories of persons who are exempt from payment, usually on the basis of income, but other reasons are used as well (see table 5-11). For instance, some countries exempt hospital staff or pensioners from payment, regardless of their income level. In addition, patients with certain conditions may be fully or partially exempted from payment, which may be justified on the basis of the principle that services subject to market failures be subsidized (treatment of communicable diseases, for example). The general rule is that no one shall be denied treatment because of an inability to pay. However, the effectiveness of fee exemptions in targeting those most in need and the extent to which the poor are cognizant of the availability of free care have not been studied. Therefore, it is uncertain whether fees limit access for the poor even in the presence of exemptions.

Table 5-11. Policies for Exemption of Persons from Payment of Hospital Fees in Selected Developing Countries

Country	Description of policy and extent of exemptions	Exemptions: Based on income	Based on age	Based on condition	For civil servants
Low income					
Ethiopia	Persons from households with income of less than 50 birr per month may apply to their local authorities for a "free care certificate" entitling them to exemption from hospital fees. The burden of proof is on the user, not the hospital. It is estimated that 30–35% of inpatients and less than 10% of outpatients in Addis Ababa hospitals and between 7 and 40% of inpatients and 20 to 60% of outpatients in rural hospitals receive free care.	X		Tuberculosis and leprosy	
Ghana	"Paupers" are entitled to free care, with such status determined by the person in charge of a health facility. This provision has been little used. Other full and partial exemptions are granted for specific conditions and procedures and for MOH staff, their spouses, and up to four children.	X		X	Health services staff
Indonesia	Indigent persons can request local authorities to certify their status and thus qualify for exemption.	X			
Mali	Exemptions exist for the poor, soldiers, primary and secondary school students, and for tuberculosis, leprosy, and psychiatric patients. Civil servants are liable for only 20% of fees; their ministries are supposed to reimburse the remaining amount (but rarely do). Nearly 70% of inpatients at one of Mali's national hospitals were statutorily exempt in 1986.	X		X	Partial

Niger	There are exemptions for indigents, students, infectious disease patients, and psychiatric patients. Civil servants and military personnel are exempt from 80% of fees; their ministries are supposed to reimburse the remaining amount but rarely do. A survey at Niamey National Hospital found that 60% of outpatients and 40% of inpatients did not pay for care.	X	Students	Infectious diseases; psychiatric	Partial for civil servants and military
Middle income					
Brazil	Government pays social insurance premium on behalf of the indigent.	X			
Dominican Republic	There are no national criteria for exemption. Ability to pay is ascertained at the hospital level by a social worker; this may involve a reduced fee rather than complete exemption. At a maximum, hospitals fully exempt 60% and reduce fees for about 50%.				
Honduras	There are no national criteria for exemption. Ability to pay is ascertained at the hospital level by a social worker. A survey at the national hospital showed that 20% of inpatients were exempted.				
Jamaica	Those qualifying for food stamps (based on a prospective means test) are statutorily exempt (between 17 and 30% of the population, assuming use of the stamps by household members), as are high-risk pregnancies. Persons with chronic conditions are partially exempted. Often, hospitals make (non-statutory) exemptions of pensioners and hospital staff.	X		Partial for chronic care	
Jordan	Persons with annual income below a defined threshold are included in a social insurance program that covers all inpatient and most outpatient costs. It is unknown how they are identified for such inclusion.	X			

(Table continues on the following page.)

Table 5-11 (continued)

Country	Description of policy and extent of exemptions	Exemptions: Based on income	Based on age	Based on condition	For civil servants
Korea	Prospective means testing for coverage of the poorest 10% by the "Medical Aid System" is implemented by local governments. Income, age, pregnancy, and immigration status are used to guide eligibility decisions. Three levels of indigence define qualifications for either free care or reduced copayments.	X	X		
Papua New Guinea	Children attending high school, adults over fifty-nine years old, and persons judged to be indigent are exempt. Children under fourteen are exempt from outpatient charges. Free services include vaccinations and injections as part of a disease control program; treatment for leprosy, TB, or sexually transmitted diseases; and blood tests for malaria or filariasis.	X	X	X	
Senegal	Broad exemptions to official fees are based on income, other patient characteristics, and medical condition. Exemption from new fees is based on certification of indigence by local government authorities.	X		X	
St. Lucia	National exemption criteria are defined based on income, student status, and participation in the social insurance program. Determination of payment status is made at the hospital level. About 93% of the population is estimated to qualify for exemption.	X	X	X	
Turkey	Fees are waived for the indigent, or are paid on their behalf by charitable institutions.	X			
Uruguay	Public hospital users are required to obtain a beneficiary card; the card is free to users below a defined poverty threshold.	X			

Former Republic of Yemen	Students and military personnel are exempt. In addition, hospital directors can exempt patients considered to be in financial hardship. An unofficial estimate is that about 25% of patients are exempted.	X	Students		Military
Zimbabwe	Persons from families with monthly incomes less than Z$150 are exempt; this criterion has not been adjusted since 1980. At provincial, district, and rural hospitals, patients can qualify on the basis of an uncontestable verbal statement (no documentation needed). In the four central hospitals, the burden of proof is on patients who must document their indigent status.	X		TB, leprosy, committed mental patients	X

Source: World Bank sources, except as noted. Donaldson and Dunlop 1987 (Ethiopia); Waddington and Enyimayew 1990 (Ghana); Prescott 1991 (Indonesia); Vogel 1988 (Mali, Senegal); Weaver, Handou, and Mohamed 1990b (Niger); McGreevey 1988 (Brazil); Lewis 1987 (Dominican Republic); Overholt 1987 (Honduras); Kutzin 1989 (Jamaica); JSI 1990 (Papua New Guinea); Russell, Gwynne, and Trisolini 1988 (St. Lucia); Marquez 1989 (Uruguay); Hecht 1992 (Zimbabwe).

Exemption policies have many distinguishing characteristics. Most are concerned with providing access to hospital care for disadvantaged groups, defined by income, age, or disease-condition. In addition, certain exemptions may have little to do with equity, often resulting in free hospital care being provided as a nonsalary benefit to civil servants. Criteria for exemption are often explicitly defined (usually in association with an official fee schedule), but in some countries there are no formal criteria. In many cases, the assessment of a patient's payment status is determined at the point of service. A few countries have prospective means tests that identify exempt persons prior to their arrival at any health facility. Countries also differ as to whether it is the responsibility of patients to prove their indigence or that of the hospital to determine their payment status.

Adhering to the principle that fee levels be consistent with ability to pay has proved to be administratively difficult to implement in practice because two of the objectives of a hospital user fee system with exemptions for the poor are in conflict: maximizing revenue collection and ensuring access for those unable to pay. The effectiveness of exemption policies in guaranteeing access has not been studied extensively, but the evidence of low fee collection rates and high percentages of persons who are exempted from fees in many countries indicates that ensuring access is generally given priority over maximizing revenue.[1] In St. Lucia, for example, the government estimates that more than 92 percent of the population is exempt under current legislation (Russell, Gwynne, and Trisolini 1988). At Hôpital du Point G in Mali in 1986, about 70 percent of inpatient days qualified for free care (Vogel 1988). In Ghana, the bulk of fee exemptions were for MOH staff and their dependents. Waddington and Enyimayew (1990) report that total revenue lost from exemptions in 1986 was estimated to be equal to about 21 percent of total collections. These examples suggest that official exemptions can reduce the level of fee collection considerably. The extent of exemptions further suggests inequitable situations in which many nonpoor persons are provided free care.

Niger's fee system includes different fees (see table 5-8) for public sector and private patients, as well as exemptions from payment based on several patient characteristics, only one of which is income. A survey of patients at Niamey National Hospital revealed that the median income of patients who had no exemptions from payment was less than the median for the entire sample of patients. The probability of paying for care for this group was higher than for several other groups with higher median incomes. There was some evidence that the poor were protected because patients in the "indigent" category were less likely to pay for care than patients in other categories. In general, however, the survey results suggest that the exemptions incorporated into Niger's

hospital fee schedule do not promote greater equity (Weaver, Handou, and Mohamed 1990b).

Some countries, particularly those with national health insurance programs, have implemented means tests to identify prospectively those who will be fully or partially exempt from hospital charges or for whom the government will subsidize an insurance premium. When they arrive at a hospital for treatment, prospectively exempt patients must generally present proof of their payment status. Prospective identification is usually performed through local government institutions and requires periodic updating of the list of those who qualify for exemption. The updating can be a costly proposition and is perhaps too much of an administrative burden for countries that are using fees to achieve only modest supplements to hospital operating budgets.

Two countries that have prospective identification of the fully or partially exempt are Korea and Jamaica. In Korea, where most hospital care is delivered by private facilities financed by a publicly mandated universal health insurance system, the poorest 10 percent in each county are judged eligible for inclusion in the Medical Aid Program, under which insurance coverage is provided free, with beneficiaries required to meet different cost-sharing obligations, depending on the indigent category to which the person is assigned (one category provides full exemption from all cost sharing, two others provide for partial exemptions). Local governments are responsible for the means testing (World Bank 1989a). Jamaica's hospital financing system is considerably different from Korea's, because its hospitals are financed almost entirely from government budget allocations, with operating budgets only modestly supplemented by revenues from user fees. Yet Jamaica also has a prospective means test that identifies individuals who will be exempted from hospital user charges. If the Ministry of Health administered this means test solely for the purpose of identifying those who would be exempt from fees, its costs would possibly be higher than the revenues raised by the fees. However, the MOH uses a preexisting means test that is administered by local government authorities to determine eligibility for a national food assistance program. Those who receive food aid also qualify for free hospital treatment. Thus, prospective identification of those exempt from hospital user fees does not impose any additional administrative costs (Kutzin 1989).

In many countries, determination of whether a patient is liable for fees is made after the patient arrives at the hospital. This may involve actually means testing the person prior to admission or treatment. Means testing at the point of service is likely to be less accurate than a formal, prospective means test, and it does place a greater burden on hospital staff and on patients seeking admission, but it is less costly than maintaining an updated exemption list. Also, in a few cases, means testing at the point

of service may be quite accurate, especially in rural facilities. This is more likely to be true in small health centers than in large urban hospitals because it is less likely that the person conducting the hospital means test is personally acquainted with the patient.

Zimbabwe does not prospectively determine patient liability for hospital fees through a national means test, but there is a national criterion for exemption. Shortly after the country gained independence in 1980, it was determined that persons from families with monthly incomes of Z$150 or less were to be exempt from user charges in government health facilities. This exemption criterion has not been adjusted despite the inflation that has occurred since that time. Presumably, therefore, as nominal incomes have increased, fewer persons qualify for exemption. The burden is on patients to demonstrate their eligibility for free care with either a pay slip or a letter from social services authorities. This requirement has not been effectively implemented, however, because few patients who claim exemptions bring the necessary documents. The result has been that, in practice, medical records or accounts clerks decide either to exempt the patient from charges or to bill the patient for services rendered. Overholt and others (1990) report that only about 20 to 30 percent of these bills are ever paid.

Examples of countries that conduct informal means tests and that do not have explicit national exemption criteria are Honduras and the Dominican Republic. In Honduras, large hospitals have social workers who develop exemption criteria that are unique to each institution. The director of each hospital has the final decision as to who will be exempt from fees (Overholt 1987). The process is similar in the Dominican Republic, where persons who indicate they cannot pay fees are interviewed by a hospital social worker. Very often when it is determined that the person cannot pay the full fee, he or she will be required to pay a lesser amount (Lewis 1987).

This review of the experience of selected countries suggests some important lessons in setting policies for hospital fee exemption. First, the purpose of exemptions should be limited to promoting equitable access to needed services and, in some cases, to increase the use of services with important externalities or which exhibit other market failures. Provision of free or nearly free care to civil servants, if this is desired by the country, should be in the form of a direct intragovernmental transfer to the Ministry of Health or the hospital. Although this policy exists in several countries, enforcement of the actual transfer has been lax. Second, national or regional criteria for exemptions should be defined and promoted so that those qualifying on the basis of age, income, or disease-condition are aware of their entitlement to care. Third, income criteria must be updated periodically or indexed to prevent inflation from limiting access for the poor. Fourth, determination of patient

payment status at the point of service appears to work adequately for countries with relatively low levels of fees, although it may be somewhat burdensome for the hospital. Patients should have to apply for exemption on the basis of the explicit criteria so that the hospital will not be burdened with the responsibility of assessing the payment status of every patient. Placing the burden for demonstrating indigent status on the patient rather than the hospital, however, may not be feasible or fair in countries with high rates of illiteracy and innumeracy.

If a country has or is planning to implement a health insurance program for its entire population, it will be necessary to establish a prospective means test to ascertain whose premium the government (central or local) must subsidize and the extent (full or partial) of the subsidy. In addition, periodic updating of the means test (including the exemption criteria) must be institutionalized to keep the system functioning accurately over time. The costs of such testing must be included in a country's analysis of the overall costs of moving toward a comprehensive national insurance program. For countries that use fees to provide a modest supplement to hospital operating budgets, the use of a prospective means test to identify those exempt from hospital fees is not likely to be cost-effective, because the cost of the means test may be a high percentage of total revenue collections. However, if there are preexisting means tests used for other purposes (to qualify for food assistance, for example), these might be a reasonable proxy for targeting the medically indigent at little marginal cost to the country.

Insurance Financing of Hospitals

As noted in chapter 4, health insurance provides financial protection against the high cost of medical treatment and unpredictable health events, and it also provides a form of earmarked savings to care for more predictable needs. The institutional arrangements for the provision of insurance differ across countries and schemes. Table 5-12 illustrates this diversity.

The extent of total population coverage is related to the overall level of employment of the urban formal sector because it is easier to enforce premium collection on this group, through deductions from wages, than on the self-employed or on others working in agriculture or the informal sector. This is reflected in countries—such as Argentina, Brazil, Jordan, Korea, Turkey, and Uruguay—that have a relatively large proportion of the labor force working in the formal sector and have significant public or quasi-public insurance programs. In many other, less-industrialized countries (for example in Ghana, Mali, and Senegal), social health insurance programs are much smaller, covering less than 5 percent of the total population. Civil servants are the group most commonly covered by

Table 5-12. Characteristics of Insurance Programs in Selected Developing Countries

Country	*Description of program*	*Population covered*	*Financing mechanisms*	*Hospital reimbursement mechanisms*
Low income				
China	All programs are public insurance: (1) Government Insurance System (GIS); (2) Labor Insurance System (LIS); (3) Communal Rural Insurance	GIS: civil servants (2.4% of population); LIS: employees of state-owned enterprises plus half of medical care costs for dependents, collective enterprise workers, and township and village enterprise workers (15.3%); Communal Rural: rural persons (10–20%). About 20% of the population is covered under major programs; another 10–15% has limited coverage.	GIS: out of MOH budget; LIS: for state enterprise workers, either entirely by employers (tax deductible) or mixed employer/employee; for collective/township/ village employees, by the employees; Communal Rural: by the participants.	With the exception of some experiments with copayments and caps, the GIS and LIS for state enterprise workers have no cost-sharing provisions for beneficiaries. LIS for collective/township/village employees and communal rural programs typically require copayments and have ceilings on total benefits.
Ghana	There is no formal insurance program, but civil servants have a type of coverage.	Civil servants (employees of government ministries).	Civil servants are reimbursed by their ministries for their hospital charges.	Does not involve the hospital.
Indonesia	All social insurance (some private firms self-insure): (1) ASKES (English translation: health insurance for civil servants); (2) Pemeliharaan Kesehatan Tenaga Kerga (health care for workers) (PKTK).	ASKES: mandatory for active and retired civil servants plus spouse and up to three dependents (about 9% of population); PKTK: voluntary pilot scheme for private sector employees.	ASKES: 2% of salary from active civil servants and of pension payments from retirees, and from interest on the program's savings deposits; PKTK: (Jakarta) 7% of wages.	ASKES: prospective per diem rate encompassing nearly all services rendered during hospitalization (still far below costs) plus diagnosis-related length-of-stay guidelines, no cost sharing unless patient opts for greater amenities;

				PKTK: similar coverage as ASKES, but rates paid to hospitals are much higher (but still below cost).
Kenya	Private: employment-based group coverage; public: National Hospital Insurance Fund (NHIF).	Private: 60,000 employees in formal sector (0.3% of population); NHIF: persons with taxable income above a defined monthly threshold and spouse and all dependent children (estimated at nearly 30% of the population).	Private: presumably from employer and employee contributions; NHIF: until mid-1990, flat monthly amount deducted at the source of income. Now, contributions are graduated to income.	Private: indemnity coverage with no catastrophic stop-loss, some with cost sharing; NHIF: inpatient services only, fixed per diem payment with defined annual maximum (no catastrophic coverage). Historically, very little has gone to public hospitals.
Mali	(1) Civil servants: reduced out-of-pocket liability for fees; (2) social security (INPS).	Civil servants; INPS: formal sector employees, same contribution also covers spouse and all dependent children. Program covers less than 1% of the population.	Civil servants: ministries pay 80% of fee out of their budget allocation; INPS—employer and employee contributions.	Civil servants: individuals pay 20% of fee, ministries liable for 80% (often goes uncollected); INPS: hospital financing uncertain, INPS provides "free" primary care through its own facilities.
Zaire	Direct insurance plan managed by the Bwamanda health zone covers hospitalization and chronic care treatment at health centers.	About 60% of the health zone's population was enrolled in 1989.	The plan has been completely financed by premium payments since it was established in 1986. Premium levels have increased annually to keep pace with inflation. Premiums are collected once a year, at harvest time, and deposited in	Case-based reimbursement, with different rates defined for each of 16 broad diagnostic categories. Each month, the hospital bills the insurance fund for services provided to beneficiaries. There is a 20% copayment, except for delivering mothers

(Table continues on the following page.)

Table 5-12 (continued)

Country	*Description of program*	*Population covered*	*Financing mechanisms*	*Hospital reimbursement mechanisms*
			a special interest-earning fund for the insurance plan. The premium collections are intended to cover hospital operating costs (excluding depreciation) for beneficiaries.	who have been given prenatal care. The plan specifies that to be reimbursable, all hospitalized cases must be referred from a health center.
Middle income				
Argentina	Social security: more than 320 autonomous health insurance institutions (*obras sociales*).	More than 70% of the population belong to an *obra social*.	The *obras sociales* are financed by payroll deductions.	There are two types of *obras sociales*: (1) direct providers of services, including hospital services, that are paid on a capitation basis; and (2) financial intermediaries, which purchase services for their affiliates from both public and private providers. Reimbursement rates are set according to a fee schedule negotiated by the *obras sociales*, professional associations, and private providers.
Brazil	Public: (1) social security institute (INAMPS); (2) civil servants of state and local governments (IPMSSSI);	IPMSSSI: civil servants (state and local) and dependents; INAMPS: active and unemployed workers and	IPMSSSI: employee and employer contributions; INAMPS: payroll deductions of beneficiaries and from general	IPMSSSI: own facilities from premiums and cost sharing; INAMPS: procedure-based prospective price paid per

	Private: (3) fee-for-service insurance; (4) employer/union facilities; (5) HMO/PPO (preferred provider organization).	dependents, active and retired civil servants of the federal government and dependents, and some indigent (90% of population); HMO/PPO: about 20 million people (who are also INAMPS beneficiaries) are enrolled in these plans (14% of population).	revenues on behalf of the indigent; fee-for-service insurance and HMO/PPO: employer and employee contributions.	case to hospitals, which are predominantly private, with no cost sharing; fee-for-service: retrospective; PPO: retrospective but with a restricted list of providers; HMO: capitated prepayment.
Jamaica	Private only, mostly fee-for-service, with one small HMO, and several self-insuring firms.	Mostly employment-based group coverage (also for spouses and dependents), plus a small number of individual plans (about 14% of population is covered).	Tax-deductible employer and employee contributions, including government in its role as an employer.	Retrospective (except HMO): public hospitals are entitled to the maximum reimbursable under an insurance plan, but they rarely receive this amount.
Jordan	All public: (1) Civil Health Insurance (CHI); (2) Military Health Insurance (MHI). It is intended to merge these into one program.	CHI: active and retired civil servants, students, poor people (income below defined threshold), and dependents of each of these (about 28% of the population); MHI: active and retired military, police, and airlines personnel and dependents (magnitude unknown).	CHI: employee contributes 2% of salaries, wages, and pensions; MHI: similar to CHI.	CHI: beneficiaries receive free inpatient and near-free outpatient care, CHI does not reimburse MOH hospitals; MHI: similar to CHI, except that contributions are channeled into military health facilities.
Korea	Compulsory national health insurance.	All.	Percentage contributions by employers, employees, and pensioners vary depending	Fee-for-service basis, with provider paid for each itemized service. Also

(Table continues on the following page.)

Table 5-12 (continued)

Country	*Description of program*	*Population covered*	*Financing mechanisms*	*Hospital reimbursement mechanisms*
			on region, employment category, and insurance society. Flat sums are paid by rural residents and self-employed, subsidized by government. Government also covers premiums of the indigent.	involves significant cost-sharing requirements.
Senegal	(1) Social Security (IPM).	Employees (and all dependents) of firms with at least 100 employees must create their own IPM, though some smaller firms may also group themselves into IPMs (the extent of total coverage is unknown).	Employers and employees each contribute 6% of employee's monthly salary, with amounts capped at a defined level.	Not much involvement with the public hospitals; there is full indemnity reimbursement with no cost sharing.
Turkey	Public: (1) civil servants; (2) Government Employees Retirement Fund (GERF); (3) Social Insurance Organization (SIO); (4) Social Insurance Agency of Merchants, Artisans, and Self-Employed (BAG-KUR). Also (5) private funds.	Civil servants and dependents (11% of population); GERF: retired civil servants and their dependents (1%); SIO: formal sector and permanent agricultural employees, retirees, military personnel, domestic servants, and all dependents (25%); BAG-KUR: members and	Civil servants: budget of each ministry; GERF: employees contribute 11% of base salaries, and government (as employer) contributes 18%; SIO: health component is 5% by employees, 6% by employers; BAG-KUR: individual contributes percentage of income up to a	Civil servants: no cost sharing, ministries reimburse MOH facility; GERF: no cost sharing, hospital bills GERF for charges; SIO: own network of facilities, no cost sharing except for drugs for dependents; BAG-KUR: comparable to SIO; private: employer reimburses facility.

		dependents (18%); private: employees of financial institutions, insurance companies, and chambers of commerce and industry (limited dependent coverage) (1%). About 53% of total population is covered in some form.	defined maximum; private: provided by employers.	
Uruguay	Social security: Sickness Insurance Fund (DISSE) contracts with health care institutions selected by each private sector worker contributing to social security. Private: prepaid organizations (IAMCs), including (1) mutual aid associations, (2) professional cooperatives, and (3) limited insurance assistance plans.	DISSE: private sector employees and dependents (12% of population); IAMCs: individual enrollees (30%) and DISSE beneficiaries. Total IAMC enrollees distributed as follows: (1) mutual aid associations—32%; (2) professional cooperatives—9%; (3) limited insurance plans—1%.	DISSE: 7% payroll tax (4% employer, 3% employee), with payments made out of this to contracted IAMC. Premium levels of the IAMC are paid on a monthly basis, are fixed by government, and are indexed. Other sources of financing are sale of services to DISSE and user charges at IAMC-operated health facilities. In addition, more than half receive state subsidies because of financial difficulties.	DISSE uses capitation payment arrangements to purchase health services from IAMCs for its beneficiaries. IAMCs own ambulatory facilities and small clinics with fewer than thirty beds. They also contract with public and private hospitals. Reimbursement modalities for physicians include salaries, fee-for-service, lump sum payments, and capitation. There are copayments for ambulatory services. IAMC members also contribute a fixed monthly amount to the National Catastrophic Health Insurance Fund to cover the costs of defined expensive inpatient procedures.

(Table continues on the following page.)

Table 5-12 (continued)

Country	*Description of program*	*Population covered*	*Financing mechanisms*	*Hospital reimbursement mechanisms*
Zimbabwe	Private: about thirty nonprofit organizations under the umbrella of the National Association of Medical Aid Societies (NAMAS). The societies are unregulated by government.	About 524,000 (5.5% of the population) public and private sector employees and their dependents. There are also some individual enrollees.	Monthly premiums are set by individual societies, with employee/employer breakdowns determined by the firms involved. Premium levels are scaled to employee income and number of dependents; 20% of medical expenses, including contributions by individual employees to medical aid, are tax deductible, and employer contributions can be deducted from taxable company profits.	NAMAS negotiates fees with private physicians for specific procedures, which are published in a relative value schedule. NAMAS also negotiates fee levels with public hospitals for medical aid beneficiaries. Although these private ward fees are higher than fees for other patients, reimbursement levels probably remain below the cost of provision; thus, government is subsidizing private care in public facilities. Copayments exist for drugs but for no other medical-hospital service.

Source: World Bank sources, except as noted. Bumgarner 1992 (China); Vogel 1988 (Ghana, Mali, Senegal); Prescott 1991 (Indonesia); Shepard, Vian, and Kleinau 1990 (Zaire); Marquez 1990 (Argentina); McGreevey 1988, Briscoe 1990, and GHAA 1985 (Brazil); Kutzin 1989 (Jamaica); Marquez 1989 (Uruguay); Hecht 1992 and Overholt and others 1990 (Zimbabwe).

publicly provided health insurance. There are exceptions to the largely urban, formal sector employee domination of the health insurance market, however. China has a large number of rural persons covered by health insurance programs, leading to moderately extensive coverage in that country. The insurance program for Zaire's Bwamanda health zone shows that it is possible to get a high proportion of participation (60 percent enrollment) in a largely rural area.

Social health insurance programs are generally financed through premium payments by employers and employees (for example, in Argentina, Brazil, China, Korea, Senegal, Turkey, and Uruguay). There are, however, no premium payments for civil servants in some countries (for example, Ghana and Mali) in which the employee's ministry is supposed to reimburse hospital charges. Reducing government health expenditure is a reason often cited for instituting health insurance, yet in many countries that have implemented health insurance programs, government remains a principal source of financing by its role as employer or through a tax code that allows for contributions to private health insurance to be tax deductible, as in Jamaica and Zimbabwe, for example.

Large-scale social insurance programs either finance hospital care directly through ownership of facilities or indirectly as third-party payers for services provided in private or public (that is, Ministry of Health) hospitals. In Turkey, the Social Insurance Organization (SIO) has its own network of hospitals, which are financed from payroll contributions without service-specific cost sharing for beneficiaries. Similar financing schemes for separate facilities are also used in many Latin American countries; for example, in Ecuador, Mexico, and Paraguay.

Two broad types of hospital payment through insurance were described in chapter 4: retrospective reimbursement and prepaid capitation. Fee-for-service retrospective reimbursement is the most common form of hospital payment found in the programs listed in table 5-12. Some programs, such as the Government Insurance System (GIS) and the Labor Insurance System (LIS)—for state enterprise workers—in China and the Government Employees Retirement Fund (GERF) in Turkey, pay 100 percent of billed charges, whereas others, such as the Korean program, require beneficiary cost sharing, typically through copayments.

Case-based reimbursement, another form of retrospective hospital payment, is used in a few countries. Recent developments in Brazil's hospital financing system through INAMPS are an example. A patient's point of entry into the health care system is through an INAMPS primary care physician. This physician determines whether hospital admission is indicated and, if so, also assigns the patient to a particular service group prior to admission. The grouping is based on the procedures to be performed on the patient, and each group has a case payment rate associated with it. This is the amount that the hospital treating this patient will be reimbursed (McGreevey 1988). Similarly, Indonesia's

insurance program for civil servants (ASKES) defines reimbursement rates according to hospital type and the amenity level for which the beneficiary qualifies. Unlike the Brazilian program, however, the rate covering a bundle of inpatient services is per diem rather than per case, and the amount reimbursed is far less than the cost of providing services (Prescott 1991).

A mix of case-based retrospective reimbursement and prepaid capitation is used for hospital payment in the insurance program run by the Bwamanda health zone in Zaire. In 1989 there were sixteen rates based on broad diagnostic classifications. Noninsured residents of the health zone are charged case-based fees for inpatient care. Insured persons, unlike the case-based reimbursement systems discussed above, are responsible for a 20 percent copayment. Another distinguishing characteristic of the program in Bwamanda is that it is direct rather than third-party insurance. The health zone manages both the hospital (and other health facilities) and the insurance plan (Shepard, Vian, and Kleinau 1990). Therefore, as in a health maintenance organization, prepayments and copayments finance the care of insured persons, rather than reimbursement of the hospital for each person treated.

Hospital financing through organized prepaid capitation plans is not widespread in developing countries, but such plans do exist in several countries, primarily in South America. In Uruguay they are perhaps furthest developed, with about 43 percent of the population enrolled in health maintenance organizations. The Sickness Insurance Fund (DISSE) contracts with private health care organizations (IAMCs) to provide ambulatory and inpatient care to beneficiaries and their dependents, making a capitation payment for each enrollee. There are generally cost-sharing requirements for ambulatory care but not hospitalization (Marquez 1989). A substantial number of prepaid capitation plans are also found in Argentina and are growing more popular in Brazil's private sector.

As expected (see the discussion in chapter 4), the use of private health insurance to finance public hospital services in developing countries is not extensive because government is implicitly providing this coverage by subsidizing public services. An illustration is provided by the pattern of private health insurance claims payments in Jamaica and Zimbabwe. Only about 2.5 percent of such payments by Jamaican insurance companies was for public hospitals (including the University Hospital) in 1986 (Kutzin 1989). In Zimbabwe in 1986–87, 3.5 percent of claims paid were for MOH services, which are predominantly those of hospitals (Hecht 1992). About 80 percent of claims payments went to private physicians and pharmacies in Jamaica, with a comparable figure of 67 percent in Zimbabwe (including dental claims). Private hospitals absorbed only 11 percent of Jamaican claims, whereas the figure for private hospitals and

nursing homes in Zimbabwe was 15 percent. This evidence suggests that in these countries in which public hospital services are heavily subsidized, the demand for private health insurance arises primarily from a desire to have financial access to private physicians and pharmaceuticals for routine random, as opposed to catastrophic, needs.

Explicit forms of health insurance that pay for hospital services exist in many developing countries but tend to cover only civil servants and the relatively small number of employees in the industrial sector of the economy. Few developing countries have national health insurance programs that have achieved, or nearly achieved, universal coverage. Movement toward national programs is most advanced in Latin America, though there appears to be growth in other parts of the world as well. Hospital financing through health insurance is least developed in sub-Saharan Africa, although the Bwamanda plan in Zaire indicates that such development is possible. Whatever the magnitude of a particular program, insurance as a means of hospital financing implies a set of incentives facing consumers and producers of hospital services. These incentives differ according to the specific reimbursement mechanisms and other institutional features in place and have implications for the efficiency and equity of the health system. They are discussed in greater detail in the next section of this chapter.

The Effect of Alternative Hospital Financing Programs on Efficiency and Equity

Alternative financing programs provide policy levers for a country to address the issues of resource mobilization and resource allocation. Changes in prices and reimbursement mechanisms, however, send new signals to producers and consumers of hospital services, which affect the allocation of health resources. In practice, there has often been conflict between the goals of resource mobilization and optimal resource allocation. Countries that have implemented relatively modest hospital user fees have not witnessed important distortions in incentives inducing changes from previous patterns of resource allocation. However, the revenues raised, though helpful, have typically been of insufficient magnitude to meet supplemental revenue objectives. Conversely, fee and insurance systems that have generated enough revenues to substitute for a significant share of what had hitherto been government budget allocations to health facilities have changed the allocation of resources, with implications for the efficiency and equity of the sector as a whole. Distortions in patterns of service use have occurred with regard to specific hospital services, between hospitals and other levels of care, and across groups of individuals (for example, insured in contrast to uninsured populations).

In table 4-3 we summarized the incentives to producers and consumers of health services generated by different types of hospital payment systems. These incentives can beneficially or perversely affect the efficiency and equity of both hospitals and the health sector more generally. In the rest of this section we present case studies of alternative financing programs from four countries—Brazil, China, Korea, and Zaire—to help illustrate the effect of various financing policies on the escalation of hospital costs, resource allocation and referral between secondary or tertiary and primary care, the use of specific medical technologies, and equity of access to health services. In addition to the system for reimbursing hospitals, the case studies describe other important institutional features of insurance programs, including the specific types of services covered by insurance, the role of the insurer in enforcing cost control, the extent of consumer cost sharing, and the extent of population coverage.

We selected the case study countries because their hospitals are financed largely, and in some cases entirely, by sources other than government budget allocations and because excellent descriptions of their health financing systems exist. Understanding how alternative financing programs have distorted the allocation of health resources and how these distortions might be mitigated is important for these countries and others considering changes to their present system of health care financing.

Brazil

Most hospital (and other personal health) care in Brazil is financed by INAMPS, the health insurance component of Brazil's payroll-financed social security system.[2] Preventive care and basic curative services for the poor are provided by the Ministry of Health and are financed through government budget allocations and administered separately from INAMPS. Social security has grown rapidly, from coverage of 23 percent of the population in 1963 to coverage of more than 90 percent in 1982. In 1983, INAMPS instituted major reforms to its system for reimbursing hospitals.

INAMPS prior to 1983. Until 1983, INAMPS reimbursed hospitals on a fee-for-service basis for each service provided to patients according to an official price list, which included 2,600 items. All costs were paid by INAMPS; covered persons were not responsible for any cost sharing. Under this system, INAMPS served as a third-party payment mechanism for services delivered by private providers. In 1981, for example, 85 percent of hospitalizations paid for by INAMPS were in private contract hospitals, whereas only about 9 percent (including 2 percent in INAMPS's own facilities) were in public sector facilities. Under the fee-for-service

reimbursement system, INAMPS served as a passive financial conduit; it did not try to impose cost consciousness on providers; indeed, they would earn more by ordering additional tests or pharmaceuticals. This reimbursement system fostered the widespread fraud (for example, billing for nonexistent patients and procedures by the private health provider network) discovered in the early 1980s. With no cost-sharing requirements, consumers similarly had no reason to be cost conscious. The private provider network "responded to a set of incentives that paid [providers] whatever they billed to provide virtually unlimited services to patients who bore none of the costs directly" (Briscoe 1990, p. 92). Largely as a result of the incentives of the curative care reimbursement system, health expenditures increased by more than 20 percent annually during the 1970s, and public sector health expenditure as a share of the gross domestic product (GDP) rose from 1 percent in 1949 to 3.7 percent in 1975 and 5.6 percent in 1982.

The fee-for-service reimbursement system encouraged hospitals to increase both the number of patients and the amount of services provided per patient as a means of generating profits. From the late 1960s until the early 1980s, the number of private hospital beds increased by over 40 percent, and the hospital industry grew faster than the rest of the economy. The curative health subsystem, financed through payroll taxes, developed independently from the Ministry of Health's public health programs and grew with increased employment and economic expansion. The MOH budget, which largely financed preventive services, did not keep pace. The result of these forces was that the share of public health expenditure devoted to curative (medical and hospital) as opposed to preventive services rose from 36 percent in 1965 to 70 percent in 1975 and 85 percent in 1982. Hospitals alone accounted for nearly 70 percent of public health expenditures in this latter year, up from less than 40 percent in the 1960s.

Because hospitals were reimbursed a specific amount for each individual service provided to a patient, they had a clear incentive to encourage the use of the most profitable treatment technologies. One such service was cesarean sections, which in 1981 accounted for 31 percent of all births, the highest rate of any country in the world. Another example is diagnostic services, particularly X rays, use of which grew at a very rapid annual rate from 1970 to 1981. Reportedly, many of these tests could have been avoided without detracting from treatment. In 1979, for example, the Rio de Janeiro State University Hospital reduced the number of X rays by 40 percent and found no loss in diagnostic efficiency. McGreevey (1988) cites this and examples of other types of complementary examinations that could have been eliminated with no effect on the treatment ultimately administered to patients. Indeed, there have likely been deleterious health consequences from such heavy use of radiological services. Overprescription of drugs, especially antibiot-

ics, was also a problem. Again, this had negative health effects, in addition to the effects of such heavy consumption on total costs.

The fee-for-service reimbursement system also skewed the allocation of public sector health resources toward costly, technologically advanced services that benefited a relatively small number of people. Among these services were renal dialysis, heart bypass operations, and CAT scans. The magnitude of this effect was such that, in 1981, total expenditure on 12,000 high-cost patients was greater than the amount spent to provide basic health services and disease control for 41 million people.

The experience of Brazil prior to 1983 highlights many issues. First, fee-for-service reimbursement of private hospitals with no beneficiary cost sharing by a public agency that exercised no control over utilization led to provider-induced increases in the volume of medical services. Second, specific services (more explicitly, technologies) with the greatest profit margin grew at the fastest rates. The combination of these factors resulted in rapid growth of expenditure on curative services and distortions away from a cost-effective mix of preventive and curative care. This distortion was especially pronounced because these services were financed and administered separately; therefore, the revenues generated for curative care through an expanding insurance system could not be used to subsidize the preventive services provided by the Ministry of Health.

INAMPS reforms. The Brazilian government introduced many reforms in its hospital reimbursement system during the 1980s in an attempt to control the growth in hospital costs and use that had arisen from the fee-for-service insurance system. A pilot program of financing and administering health services was implemented in one relatively prosperous state. This program, called the Curitiba Plan, had features aimed at many of the problems inherent in the existing system. Prior to implementation of the Curitiba Plan, for example, private hospitals treated many walk-in patients as emergency cases and thereby received substantially higher fee-for-service reimbursements from INAMPS. The Curitiba Plan created a structured referral system that gave public primary care providers ("physician-auditors") authority to set the course of treatment for all patients and diminished the financial incentive to recommend that patients be hospitalized in private facilities. Under the initial Curitiba project, hospitalizations were reduced to 5 percent of initial consultations, as compared with a Brazilian average of 6.5 percent. By mid-1982, INAMPS declared that the plan had reduced hospital admissions by 30 percent.

Another feature of the Curitiba Plan was a case-based retrospective hospital reimbursement system. This financing mechanism gave clear incentives to providers to economize on the treatment provided to an

individual patient (see discussion of case-based reimbursement in chapter 4). For example, a private hospital would not be reimbursed for ordering tests in addition to those prescribed by the physician-auditor, and such a hospital would also have a clear incentive to limit each patient's length of stay. Despite these changes, however, there are still no cost-sharing provisions for INAMPS beneficiaries, so patients have no financial incentive to limit their demand for medical services.

These features of the Curitiba Plan were implemented more broadly throughout INAMPS in 1983. By requiring entrance to services to be through public primary care providers (that is, its own employees) who had the responsibility for defining a patient's course of treatment for which prices were already set, INAMPS took a more active role in enforcing cost discipline on private providers. The reforms did not alter all the problematic features of the health insurance system, however. With a fixed reimbursement rate per case type, hospitals could maximize profits per case by minimizing the level of inputs per admission, which raises concerns about the quality of care. Although a case-based reimbursement system encourages providers to economize on the level of resources allocated to the care of an individual patient, such a system also provides incentives for hospitals to maximize the volume of profitable admissions. Because there are a limited number of payment categories, there will be patients of varying severity within each category. There is some evidence that this case-based reimbursement system led to a change in hospital case mix. Rodrigues (1989b) concluded that private hospitals shifted more costly patients to public and university hospitals, leaving themselves with a mix of less severely ill patients for which profit margins per patient were greater. Such shifting could lead to intense pressure on public facilities.

A reform of the Brazilian health care financing system, announced in 1987, is the Unified and Decentralized Health System (SUDS), which may address the problems of escalating costs of curative care and distortions in the allocation of resources across preventive and curative interventions. By integrating management of all levels of care and of contracts with suppliers of private medical services at the state level, SUDS establishes an administrative framework that is consistent with an effective referral system. In addition, SUDS unifies the functions and resources of INAMPS and the MOH at the state and local level. This strategy allows for cross-subsidization of preventive (MOH) services from the revenues generated through the insurance system (INAMPS), although there are as yet no data from which an evaluation can be made. Such a reallocation of health resources could lead to a more cost-effective mix of preventive and curative care services.

Brazil's attempts to address the problems of its health care financing system appear to generate better incentives for a more efficient health care delivery system, although the available data are insufficient to

evaluate their effect. The reforms that began under the Curitiba Plan created an institutional framework to guide consumers through the health care system. Another aim was to limit provider-induced demand by requiring that the gatekeepers to the system be employees of the insurer and that these persons determine the total course of treatment to be provided. Fixed reimbursement on a case rather than fee-for-service basis should reduce the number of unnecessary procedures that are performed. Finally, unifying the resources and administration of the institutions responsible for delivering both curative and preventive services allows for cross-subsidization and the possibility of a more cost-effective health care delivery system.

China

China's system of health financing has two characteristics that are crucial to the incentives it creates.[3] First, financing of hospitals and other health institutions was decentralized in 1981, and these facilities were instructed to cover their nonstaff recurrent costs through user charges. The second characteristic is health insurance programs, most of which reimburse hospitals (and other providers) 100 percent of charges. Overall, about 20 percent of the population is covered under such programs, and another 10 to 15 percent of the population has more limited insurance coverage (that is, with cost-sharing provisions). Reimbursement under these programs is on a fee-for-service basis, and insurers are third-party payers who are separate from the providers of care and who play no role other than financing services. Provision of more services, meeting the demand for services, and the decentralized financing incentives are linked to the expansion of capacity rather than to increased efficiency in the provision of services.

The decentralization reforms, which were intended to reduce the health sector's drain on public budgets by requiring most health institutions to earn revenues sufficient to cover operational costs and some capital costs, have resulted in supplier-induced consumption of those health services on which a profit (that is, a positive margin of revenues over costs) can be earned. The pricing of hospital services, coupled with the knowledge that most insurance will reimburse 100 percent of charges billed to their beneficiaries, has encouraged health care providers to offer more and higher priced services. Rapid escalation of health costs has been one result of these policies. Total health costs have grown steadily from 2.6 percent of GDP in 1980 to 3.2 percent in 1988, and total health expenditures grew at an annual average rate of 17 percent in real terms between 1980 and 1988. Most of this growth has resulted from increases in fee payments by patients and in insurance reimbursements.

The pricing of hospital services has created distortions in the mix of services provided. Fees for technologies and services that were in use in

the 1950s (for example, outpatient consultation, inpatient care, surgical operations) were set well below cost at that time, and even with price reform in 1985, fees for most of these "older" medical technologies, procedures, and services are considerably below cost. Older price guidelines do not exist for new technologies, and case studies of the service costs and prices for many of these (for example, electroencephalography, coronary care, CAT scans, ultrasound, and renal dialysis) reveal strong financial incentives for their provision. There is an incentive for hospitals to use equipment frequently, especially for those insured patients who bear none of the cost of these procedures. To the extent that prices charged for the use of these technologies are greater than their marginal cost to the hospital, the hospital makes "profits" that can be used for yet more sophisticated equipment purchases and expansion of services, paid into workers' welfare funds, or given as incentives and bonuses.

Another category of hospital prices is drugs. Hospitals are explicitly allowed to mark up the price of drugs by 10–18 percent above the wholesale price for which they acquire them. While the markup is intended to cover storage and distribution costs, there are clear incentives to overprescribe. Overprescription is probably occurring, as evidenced by an average of 2.3 drugs prescribed per patient contact and the nearly 50 percent of total health expenditures going for drugs. As in Brazil, this is not just a financial problem; overprescription is clearly contrary to sound medical practice and can have deleterious health effects. Pricing policy has led to a de facto subsidy for the domestic pharmaceutical industry by the health system, and particularly the insurance systems.

Another consequence of the incentives of the financial decentralization reforms in China is that government health expenditure, and to a much greater extent total health expenditure, has been increasingly devoted to higher-level hospitals at the expense of lower-level hospitals. Data on government health expenditure show that the share of higher-level hospitals rose from 38 percent in 1980 to 44 percent in 1987, whereas that of basic hospitals fell from 23 to 17 percent during the same period. Antiepidemic and maternal-child health activities remained about 17 percent of public expenditure. The growth in the percentage of total health expenditure financed by fees and insurance (from 72 percent in 1980 to 81 percent in 1987), however, and the preponderance of fee and insurance payments devoted to hospital services, meant that the share of total health expenditure devoted to basic and preventive services undoubtedly fell during the period.

Changes in the share of total health expenditure between the insured and uninsured during the 1980s reflect a skewing of health service provision away from the uninsured majority of the population. By 1988, insurance reimbursements accounted for 44 percent of total health financing, yet only about 20 percent of the population was covered under

official insurance programs in that year. On a per capita basis, consumption of health care by insured persons is much greater (and is rising dramatically) than by uninsured persons, suggesting some moral hazard on the part of the insured and possibly limited access for low-income uninsured persons.

As in Brazil, the experience in China suggests that third-party fee-for-service reimbursement of 100 percent of hospital charges leads to rapid cost escalation. In a system in which only part of the population has insurance coverage, there is clear evidence of differential use of services, although the relative extent of moral hazard on the part of the insured and lack of access for the uninsured is uncertain. China's experience also highlights the importance of price incentives for the use of particular medical technologies. The signals sent to providers have resulted in the expansion of profitable high-technology services and drugs, which is probably not a cost-effective use of hospitals. In addition, because investment decisions are based on existing signals, the long-run consequences of today's distorted price signals are likely to be an even more inefficient mix of facilities and equipment. In short, although China's decentralized pricing system and insurance programs have been effective in generating revenues for hospitals, they have done so at the cost of equity in the use of services and inefficiency in their organization and production.

Republic of Korea

Korea introduced a compulsory social security health insurance scheme in 1977 that initially covered less than 15 percent of the population.[4] Since that time, legislation has rapidly increased the beneficiary population, and universal coverage was achieved during 1989. Insurance is provided through a large number of nonprofit, noncompeting societies organized at the firm, firm group, or county level. There are programs for the unemployed and self-employed, plus public assistance programs that subsidize the insurance premium for indigent persons (about 10 percent of the population). A national fee schedule is negotiated annually, and hospital payment rates are set on a cost-plus basis. Providers are retrospectively reimbursed on a fee-for-service basis according to this schedule. Unlike the program in Brazil and most health insurance in China, there are significant cost-sharing requirements in the Korean program, except for certain categories of the medically indigent. A combination of deductibles and copayments yields effective coinsurance rates of 62, 65, and 41 percent for outpatient services in upper-level hospitals, other hospitals, and clinics, respectively. The inpatient coinsurance rate is 20 percent at all facilities. Extra payments to senior physicians and other undocumented expenses, however, mean that the extent of actual cost sharing is even greater.

The combination of third-party benefit payments and cost sharing serves to finance hospital care, which is largely provided by private facilities. The insurance societies and the government exert more control over providers than do their counterparts in China (and in Brazil prior to 1983). In mid-1989, for example, a law was passed requiring patients to have a referral slip from clinics before being allowed to use hospital care. The Ministry of Health also defines the services that are covered by insurance and ultimately determines the prices for covered services. A number of expensive, high-technology services, such as CAT scans and extracorporeal shock wave lithotripsy (ESWL), are not covered, however, and providers can charge unregulated prices for these.

Attempts to sort out the effects of the expansion of insurance coverage on total Korean health expenditures are confounded by the effects of the rapid growth in per capita income (about 8 percent annually during the last two decades) on demand. Total health spending grew from 2.7 percent of GNP in 1976 to 5.1 percent in 1985. This growth resulted from a combination of increases in cost per case (inpatient cost per case nearly tripled and outpatient costs nearly doubled from 1980 to 1988, as compared with an increase of only 60 percent in the consumer price index over the same period) and in the volume of services used (physician visits per capita nearly doubled and hospital admission rates rose by about 40 percent between 1980 and 1988). The health insurance system magnified the increased utilization and costs that would have been expected as a result of income growth and an aging population because the fee-for-service system using cost-plus pricing rewarded cost-increasing rather than efficient behavior. Despite the presence of significant copayments in the system, hospital utilization data indicate a growing tendency to increase the volume of services provided; the frequency of repeat visits is rising and lengths of hospital stay are increasing. These increases suggest that incentives on the demand side (that is, substantial cost sharing) have not been very effective at limiting service use in the face of rapidly rising incomes.

The insured population shows a marked tendency to use large general hospitals even for primary care; these hospitals tend to have higher proportions of specialists and advanced medical equipment and are accordingly more costly. Using general hospitals is also the most expensive option for a consumer: such hospitals are allowed to charge fees that are about 15 percent higher than they are at clinics, and the percentage copayment for outpatient care is 50 percent greater. Despite these higher charges, however, the use of general hospitals has risen at almost the same rate as that of clinics in recent years. The pattern of overuse of the more sophisticated general hospitals that is becoming established in Korea may be incorrectly interpreted by planners as a need for more of these facilities, which could generate still more undesirable cost escala-

tion. The referral requirement introduced in mid-1989 appears to be an appropriate device for initially steering patients toward lower-cost providers, but there are as yet no data to suggest how effective this has been in practice.

Because prices are set by the government, Korean hospitals must compete for patients on some other basis than price. This other basis primarily has been the availability of senior medical staff and of sophisticated medical equipment. Given the strong societal preference for senior physicians and new technology, it is not surprising that hospitals are acquiring more and more modern equipment. For example, CAT scanners are now in nearly every hospital with more than two hundred beds. The costs of these devices generally are greater than the revenues collected from their use, but hospitals use them as loss leaders to attract patients. Other technologies used in this manner are heart transplant facilities and ESWLs. Korea has thirty-eight heart transplant centers and twenty-six ESWLs. Canada, by way of comparison, has only thirty-two and four, respectively.

Universal coverage under Korea's national health insurance program theoretically should provide for universal access and better equity than had previously existed. However, the very high coinsurance rates charged to all patients irrespective of income (with the exception of some categories of the medically indigent) raise equity concerns because they are more of a burden for the poor. Data on per capita health service use by income class for insured persons in 1987 show that use by the poorer insured was considerably less than that by higher-income insured persons. In addition, the capacity to make informal "under the table" payments and to pay for noncovered high-technology services has reportedly led to a two-class system of health care.

Despite significant beneficiary cost-sharing requirements in Korea's fee-for-service hospital reimbursement system, the introduction and expansion of insurance has been associated with rapid growth in health and hospital costs. In addition, demand-side strategies designed to encourage appropriate use of referral facilities have not been very effective. This suggests that the key to encouraging a more efficient service delivery system is to create strong incentives to providers to limit the volume of services they provide and use more cost-effective treatment settings. Clearly, such incentives do not arise from Korea's cost-plus fee-for-service system, although the referral requirement added to the system in 1989 should mitigate some problems if implemented forcefully. Korea's insurance program uses its strength as the principal purchaser of services to negotiate a national hospital fee schedule, but an unintended consequence of this has been the means by which private hospitals compete for patients on the basis of perceived quality and

status. Because consumers' understanding of hospital and medical practice is usually limited, there may often be great differences between perceived quality and actual quality. As a result, hospitals have competed by accumulating technologically advanced equipment to attract patients. As in China, this action creates a distorted pattern of investment that will lead to an increasingly inefficient mix of services in the future. Finally, equity problems remain despite the achievement of universal coverage. The heavy official cost-sharing obligations and the unofficial payments demanded by private providers weigh most heavily on lower-income insured persons, who consequently use fewer health services than the rest of the population.

Zaire

The focus of this case study is on only one small region of Zaire, not a national system of hospital financing as was the case in the other examples.[5] In 1975 the government developed a plan to organize its health services around a large number of "health zones," each serving approximately 100,000 people with a reference hospital and satellite health centers. The health zones were given a considerable degree of autonomy, including responsibility to develop cost-recovery programs to meet their operating and maintenance costs (Bitran and others 1987). In 1985 the management of one of the zones, Bwamanda, created an insurance program as a means of generating revenue for its reference hospital and organizing the delivery of services in the zone. Unlike the insurance programs discussed earlier, the Bwamanda insurance plan covers only hospital services (plus chronic care treatment in health centers). In addition, it is organized and managed by the health zone and is thus a direct rather than a third-party program, and it is administered as a prepaid capitation health care organization. Enrollment in the plan is voluntary, but to limit the effects of adverse selection, all members of a family are required to join if one member joins. For residents of the zone who are not enrolled in the plan, case-based fees are charged, with patients assigned to one of sixteen categories according to their clinical characteristics. The insured population pays 20 percent of the case price charged to uninsured residents of the zone. In addition, the hospital charges nonresidents twice the resident rate and employed persons (whose employers are required by Zairian law to pay the full cost of their medical care) 250 percent of this rate.

Even though the insurance plan does not cover charges in health centers (except for care of chronic conditions), it includes strong incentives to discourage self-referral to the hospital. Specifically, the plan will pay for hospital care only if a beneficiary has a referral slip from a health

center. The incentive for first using less-complex facilities is thus built directly into its reimbursement rules. Another aspect of the plan that encourages appropriate use of primary care facilities is that insured women who have received prenatal care are entitled to a free hospital delivery.

The uninsured were found to have hospital admission rates that were lower than for those insured through the plan, and much lower (nearly twentyfold) than for those whose fees were paid by their employers. This suggests a combination of limited access for the very poor, because there are apparently no exemptions from payment, and moral hazard arising from the financial protection provided by insurance and employer payment. The enrollment in the plan of individuals who believed they were likely to use the hospital (adverse selection) was probably also a factor in the overrepresentation of plan members in the use of hospital services.

The purpose of establishing a hospital insurance program in Bwamanda was to increase the level of resources available for health services. Increases have been achieved, yet there have not been gross distortions created in the process, as has occurred in other countries. There is no evidence to suggest that the case-based reimbursement and pricing system has resulted in distorted incentives for the use of any particular technologies. The direct insurance model, wherein the insurer and provider are the same entity, offers several advantages over third-party payment in regard to the efficient organization of the health care system. For example, no specific referral incentives to providers are needed; therefore, referral rules are more easily enforced. The principal drawback of the voluntary insurance plan used in Bwamanda is that access to care is not the same for the uninsured and the insured (including employer-paid). The greater use of services by the insured reflects moral hazard arising from insurance coverage; it may also mean that access for the uninsured is limited by their unwillingness or inability to pay. The latter is of greater concern, but the magnitude of this problem is unknown. One possible way to encourage greater equity would be to subsidize insurance premiums, fully or partially, for those unable to pay, although this would involve means testing that might be relatively costly to apply. Another issue related to the Bwamanda plan is that it has a monopoly on the provision of hospital services in the region. This may limit the replicability of this model in other regions unless they have the resources and the will to ensure maintenance of the quality of services. Overall, however, the Bwamanda health plan has successfully organized and sustained the services provided in the zone, demonstrating that insurance can be a viable option for financing health services in a rural area with a population primarily composed of self-employed farmers.

Synthesis

These case studies suggest that the role played by the institution administering an insurance program is perhaps the principal determinant of the effects of the program on the efficiency and equity of health resource use. In particular, insurers should not merely channel funds to providers; rather, they should take a very active role in establishing mechanisms (such as contractual obligations) that encourage providers of health services to make efficient and equitable resource allocation decisions. By controlling supply and strongly influencing demand for services, providers can increase the use of curative care services in a manner that, as the examples of Brazil, China, and Korea show, leads to rapid escalation in health care costs, distortion of resource allocation, inappropriate use of medical technologies, and inequitable access to the services available.

Cost escalation. In each of these countries, alternative financing mechanisms were introduced as a means of increasing the level of resources available in the health sector and reducing government subsidies. Thus, it is not surprising that each has been associated with rapid growth in expenditures for health and hospital care. Cost escalation becomes problematic when the incentives of a specific financing mechanism cause expenditures to spiral out of control. These studies indicate that third-party fee-for-service reimbursement leads to large and rapid increases in hospital costs because of provider incentives to increase the volume and even the cost of services. The presence of significant cost-sharing requirements in Korea apparently did little to mitigate this cost escalation, although more analysis is needed to isolate the partially offsetting effects of rapid income growth and cost sharing on service demand. Certainly, however, providers have a very important effect on the quantity of services ultimately demanded by health consumers, given the ignorance of most consumers of their medical treatment needs and options. Greater cost containment can be achieved by creating incentives for providers to limit the volume of services rather than focusing efforts solely on the demand side through price signals.

Case-based reimbursement should result in less cost escalation than a fee-for-service system, but the evidence from Brazil is insufficient for an evaluation. Case-based reimbursement does not eliminate the incentive to increase costs; it does motivate health care providers to maximize profitable admissions and minimize the quantity of services provided per admission. This incentive is likely to limit cost increases somewhat. It may, however, have a negative effect on the quality of care, and either competition or regulation are necessary to maintain standards. Furthermore, where a mix of private and public providers are financed on this

basis, as in Brazil, private hospitals have an incentive to alter their case mix by taking less severe, more profitable patients and leaving sicker patients for the public hospitals. This can result in an overly strained public system.

Resource allocation and referral. Because third-party fee-for-service reimbursement by a passive insurer generates rapidly rising costs for hospitals, and, more generally, for curative services, it tends to lead to a decline in the relative share of health resources devoted to preventive services. This does not necessarily imply a decline in resource availability for prevention, but for most countries the marginal benefit of additional resources in preventive services would be greater than in curative services. This effect can be mitigated if the provision of preventive services can be subsidized from the surplus generated by the expansion of curative care. Such cross-subsidization is most feasible when a unitary authority is responsible for all levels of health care, from primary prevention to hospital services. Two models for this are the direct insurance plan used in Bwamanda, in which the health zone is responsible for all levels of care but generates revenues through the provision of curative care, and the SUDS in Brazil, which unifies the resources and administration of curative and preventive services. Both strategies should be closely monitored to determine how effectively they allocate resources.

Financing systems can generate incentives for the use of specific types of facilities by price incentives to consumers or by fiat. Evidence from Korea suggests that price incentives alone may not work very well if perceived quality differences between levels of facilities are great. Making reimbursement contingent upon a patient's use of specified points of entry into the provider network appears to be more effective. In Zaire's Bwamanda health zone, this functions as a price incentive, with a very large (80 percent) penalty for insured persons who self-refer to the hospital. Brazil's Curitiba Plan suggests a strategy for countries with a diverse mix of providers: require entry into the network to occur through a primary care facility run by the insurer (an aspect of the direct insurance strategy). This requirement not only reduces self-referral to hospitals, it also removes the decision on the course of treatment from the provider who will ultimately be reimbursed.

Incentives for the use of medical technologies. The financial incentives that influence the use of a medical technology arise from how these specific services are priced, as is evident from the experience of Brazil and China. Hospitals have an incentive to maximize use of those services for which the margin of reimbursement above the cost of provision is greatest. Consumers rely on the medical judgment of providers and often tend to equate new, high-technology procedures and a greater quantity of phar-

maceuticals with better health care. When the price or reimbursement rate differs from the marginal cost of a service, the incentive may shift the provision of services away from what is medically appropriate by introducing financial considerations into the decision about the choice of technology. The consequences of this distortion in regard to long-term efficiency can be severe, because investment in future capacity is often based on current levels of service use. The problems can be mitigated to some extent by using case-based rather than retrospective reimbursement, but overprovision incentives may remain, depending on the nature of the case payment categories.

Equity of access. Insurance coverage is a means of ensuring access to medical services that may otherwise be difficult for many persons if prices for services are high enough to recover costs. When coverage is universal, as is practically the case in Brazil and has become true in Korea, differences in access arising from insurance status should not occur. When only part of the population is covered by insurance, as in China and Zaire's Bwamanda health zone, disparities are likely, because the cost to the uninsured of using health services is radically different from the cost to the insured population. Such disparity does not necessarily mean that access for the uninsured is inadequate, however. Although heavier utilization among the insured population (moral hazard) can be expected, it is not a serious problem so long as sufficient access for the uninsured can be guaranteed. Access problems may exist even in the presence of universal coverage, however. Briscoe (1990) shows that in Brazil, for example, poor rural areas may lack sufficient facilities or trained personnel even when the population has insurance coverage. In Korea, access is limited for poor persons who cannot make informal cash payments to providers.

In poor countries in which lack of access is clearly a problem, the issue of publicly subsidized insurance coverage for one segment of the population raises serious questions of equity. Partial insurance for a relatively well off section of the population exists in a number of countries, even those whose health services are priced modestly and are thus heavily subsidized by government. In Indonesia, for example, beneficiaries of the compulsory health insurance program for active and retired civil servants (ASKES) are entitled to free use of public hospitals. The hospitalization rate of the beneficiaries of ASKES is about five times the national average, and most ASKES families are probably in the top quartile of the Indonesian income distribution. Their heavy use of services, which remain heavily subsidized, implies that this relatively well off segment of the population captures a disproportionate share of public subsidies for health as a result of the easier access they have because of their insurance coverage (Prescott 1991). This issue may also be of concern in

several African countries (see Vogel 1988 for brief descriptions of health insurance in Côte d'Ivoire, Mali, and Senegal).

Summary and Recommendations

The use of fees and risk sharing offer the possibility of reducing the financial dependence of the hospital system on public sector funds while meeting equity objectives, but this possibility has not been exploited in many countries. The general conclusion from the review of revenue experience in this chapter is that in most countries revenues generated from nongovernment alternative financing sources remain a small fraction of total hospital expenditures. Surprisingly, the percentage of costs recovered does not appear to be related to the level of per capita income or the prevailing ideology of a country. The experience with cost recovery in such diverse countries as Brazil, China, Turkey, and Zaire demonstrates both its feasibility and usefulness as an alternative source of revenue for hospitals.

Many countries have formal fee policies for hospitals but collect only negligible revenues. In these countries, the low level of unit fees, loose enforcement of payment of these fees, and broadly defined exemptions contribute to the small size of collected revenues. Commonly, the failure to adjust for inflation erodes initially modest fees to the point that they become negligible. Inadequate enforcement adds to the problem and results in the collection of substantially less money in fees than could be collected, even when the existing fee structure is used as a basis for calculating the potential.

The usefulness of fees to achieve the objectives of equity and efficiency is impeded by the lack of a systematic relation between fees for specific services and the related marginal or average costs. Ideally, to promote efficiency, fees should act as a device to encourage the rational use of resources, but fees kept at negligible levels cannot provide the correct incentives to achieve efficient resource use. Also, to promote equity, fees for services used by higher-income households can be set at levels that return a profit to subsidize more broadly used services. More often, however, fees are set well below costs even for private amenity hospital services, with the inequitable result that such services are subsidized from general government revenues.

Review of the experience with fees suggests some practical additions to the guidelines for fee policy set out in chapter 4:

- The routine adjustment of fees over time should be an institutionalized administrative action rather than a political decision. This can be done by tying fee levels to price indexes.

- The structure of fees should be simple. The price structure should be the least detailed in situations in which the cost-recovery goals are modest, but in any case a simple structure is easiest to administer. The complexity of fee structure is also limited by the degree of institutional and management development.
- Price setting should be decentralized in large countries and in countries that have great geographic variation in per capita income.
- Some degree of fee retention should be allowed so that hospitals are provided with a financial incentive to collect fees. These revenues should be used by the hospital to increase the quality of services and thus increase the willingness to pay for hospital care.
- Exemptions should be limited largely to those made on equity grounds. Extra subsidization and often complete exemption are also warranted for services that have substantial benefits well beyond those to the patient, such as treatment of communicable diseases.

As long as fees remain low, allowing for the recovery of, say, less than 30 percent of total recurrent hospital costs, exemption policy and government subsidies can absorb the financial risk of high-cost illness. A higher degree of cost recovery, however, requires some form of explicit risk protection. Whether the form of hospital payment is retrospective reimbursement or prospective capitation payments, the presence of high user charges and insurance affects the efficiency and equity of the system. In many countries, insurance and fees have led to cost escalation in hospitals caused by incentives for providers and consumers to produce and use services and for providers to adopt cost-ineffective technology. Retrospective reimbursement can lead to a declining share of health resources used for prevention because of the rapid expansion induced in the provision of curative care. Government-subsidized insurance plans that cover only a small percentage of the population create the possibility of great inequities in the provision of and access to services.

The analysis in chapter 4 and the experience of selected countries reviewed in this chapter lead to some recommendations for the design of a risk-sharing system to avoid these problems.

- The institution administering the insurance program should take a very active role in establishing mechanisms that encourage providers of health services to behave in a manner consistent with the goals of efficiency and equity. This is most easily achieved with direct insurance, but it is also possible for third-party payers.
- Cost sharing is needed in either reimbursement or prepaid capitation schemes to provide consumer incentives for efficient use of ser-

vices. More important, however, are mechanisms to change provider incentives, because cost sharing by itself is unlikely to be sufficient for effective cost containment.

- In countries with large social insurance programs, revenues generated by the expanding curative care sector should be used to cross-subsidize preventive services. Such a program is facilitated when all levels of care are managed by a unitary authority on a geographic basis.
- Prepaid capitation and direct insurance schemes are recommended over retrospective reimbursement as practical means of cost recovery and provision of managed care in circumscribed communities.
- Case-based reimbursement is preferred to a fee-for-service system to diminish provider incentives to overproduce.
- In most cases, publicly sponsored insurance for only a small part of the population should be avoided. The insured provision of subsidized services will add to the inequity of resource use across income groups. If sufficient political will can be mobilized, however, profits can be generated from the provision of insured services to a small population group. Then insurance can allow cross-subsidization, but the situation can quickly erode because of the greater political power of the high-income insured.

In considering changes to existing health financing policies, some contrasts can be drawn between the elements of alternative policies that might be suggested for low- and middle-income countries. The elements of a recommended hospital cost-recovery strategy can be differentiated by the levels of per capita income and institutional development for a prototypical country. Individual countries, such as China, that have low per capita income yet advanced institutional development provide exceptions to this model. The evolution of hospital fee policies from low- to middle-income developing countries is tied to information and recordkeeping capacity, a concern with equity, the potential for insurance, and the quality of services.

In countries with low income and poorly developed institutions to administer fee collection and provide complementary management information, equity requires that fees be kept low, and the difficulty of administering fees requires a simple fee structure. With low fees and poorly developed institutions, reimbursement insurance may not be practical or required. The experience reviewed above for selected low-income countries does suggest, however, that prepaid capitation schemes can be practical in circumscribed regions in which community support and institutional development provide the needed environment. The quality of services affects the success of cost-recovery measures, and, likewise, if a portion of cost-recovery revenue is retained for the use of the hospital, these measures can improve quality.

For middle- and high-income countries, fees collected can be more substantial. A higher level of fees should be accompanied by development of the institutional infrastructure to manage fee collection, recordkeeping, and data collection. Some increase in the complexity of the fee structure may become practical if the objective is to achieve greater equity and efficiency through cross-subsidization and to make fees reflect costs more closely. Higher fees will also make it more practical to cover the cost of collection and enforcement of exemptions. Some means of risk sharing must be introduced if fees are increased, and the use of reimbursement insurance will also put greater demands on data collection and recordkeeping. Higher fees, too, will mean that even greater attention must be paid to the quality of services.

For both high- and low-income countries, the level and structure of hospital fees should reflect the role of the hospital in the referral system. Fees can act as a market signal to direct client demand to the most efficient service providers. This aspect of fees is considered further in the next chapter.

Notes

1. Waddington and Enyimayew (1990) report anecdotal evidence that suggests that in Ghana, unlike in many other countries, there is a reluctance to use the exemption based on income. As a result, some people are foregoing treatment because of the cost.

2. Information on Brazil is drawn from four sources: Briscoe 1990; World Bank 1988; McGreevey 1988a; and Rodrigues 1989b.

3. Information on health and hospital financing in China is based on Bumgarner (1992).

4. Information on hospital and health financing in Korea is derived largely from Yang 1991, and World Bank 1989a.

5. Information on the financing of hospital services in Zaire is derived from Shepard, Vian, and Kleinau 1990.

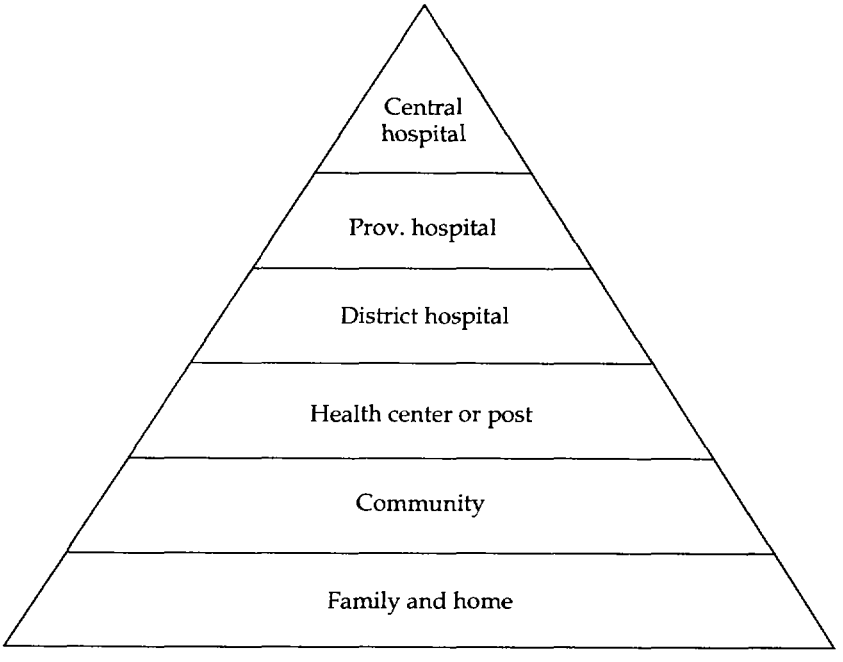
6. Reallocating Hospital Resources to Improve Health Services

This chapter covers the role of hospitals in the health service system and suggests changes that can improve the use of health sector resources. Hospitals are an essential part of the health system, but the reallocation of resources to nonhospital alternatives can reduce the cost of providing health services. Such reduction does not necessarily imply budgetary savings; rather, it may imply increasing the number of people who are served or improving the quality of services for a given level of expenditure.

The term "hospital" connotes an image of a physical building in which services are delivered by a skilled staff. Traditionally, hospitals have provided a focus for the delivery of interventions requiring special personnel skills and equipment, monitoring of patients, or containment of patients for therapeutic reasons. Most health care, however, actually takes place outside of hospitals, in clinics, medical offices, pharmacies, schools, and homes. Hospitals, by providing a technical focal point for the referred delivery of skilled care, can magnify the effectiveness of nonhospital health care.

One common description of the health sector uses a pyramid, such as that in figure 6-1, to emphasize that hospitals are the end point in a referral hierarchy that provides most care at the bottom of the pyramid through primary health care. Although the pyramid is a stylized depiction of a health system and does not illustrate the relations between and the functions of the different levels of the system, it does clearly describe the principle of providing most care at the lower levels. Movement between levels of the pyramid occurs, theoretically, by referral and according to need. Within this general view of the health system, however, there is considerable flexibility to redefine specific functions carried out at different levels in the system and to change referral patterns.

Figure 6-1. Health Referral System



The pyramid is not merely descriptive. It is prescriptive, suggesting how the health care system should be organized. There is an economic logic inherent in the structure of the pyramid, which reflects the relative cost-effectiveness of alternative technologies for providing care. The width of the pyramid at any point represents the relative number of persons who should be served by a particular provider category. The criteria for determining which provider should provide care at a given level are (a) epidemiology (the frequency of various diseases and conditions in the population) and (b) the unit costs of providing services at each level. At the bottom of the pyramid should be the most frequently occurring conditions that are the least expensive to treat, whereas at the top should be the rarest conditions that are the most expensive to treat. This depiction is consistent with the goal of patients to minimize the costs (including travel and time) of having their illnesses treated.

Each provider category represents, in broad terms, a technological alternative for treating diseases and conditions arising in the population. Health planners can use epidemiological analysis, a knowledge of technological alternatives, and the unit costs of these alternatives to derive a cost-effective network of providers for their country. Mwabu (1989)

notes that cost-effective delivery of basic health services occurs within the context of a particular organizational structure of facilities. Because different structures, or delivery systems, have different costs of delivering the same package of basic health services, the structure with the lowest cost (that is, the one that would maximize population coverage, given resource constraints), is the most cost-effective one.

For each provider category, health care elements such as use of pharmaceuticals, treatment, disease control, maternal and child health and family planning (MCH/FP), water and sanitation, nutrition, and health education are relevant. Providing these elements of health care involves functions such as information, management, personnel development, logistics, facilities, and research. There are linkages between health care elements and functions at all levels of the health care system. This conception of the health care system recognizes, for instance, that the household and community can provide labor, facilities, logistics, and other functions for treatment and disease control. Or, to provide a contrasting example, all levels of hospitals can be involved in nutrition, health education, and basic MCH/FP. Simply the recognition that specific services can be provided at different levels of the system does not imply that these services should be provided at many levels (or at least subsidized to the same extent across all provider categories). Nevertheless, this concept recognizes that there is considerable interaction between the levels with regard to all the functions and health care elements (WHO 1987a).

The growing use of alternatives to the traditional structure of hospitals represents an effort to increase the delivery and improve the effectiveness of health care services in low-cost settings. This notion is consistent with efforts to encourage patients to use lower levels of the referral pyramid. A more cost-effective network of providers can use hospital resources to support primary health care and create new institutional settings for care at the community and family levels. Because we are concentrating on hospitals in this book, we make no attempt to provide a complete model of the interactions between all providers in a referral network (see King 1966 for a detailed discussion). The focus in this chapter is on the role of the hospital, including the potential for hospitals to support primary health care, the reallocation of patients and services across basic and tertiary hospitals, and the use of innovative alternatives for providing care ordinarily delivered on an inpatient basis in hospitals. Accordingly, in the first section of this chapter we discuss possibilities for improving the overall cost-effectiveness of hospitals by changing referral patterns and thus the mix of services delivered in lower-level hospitals and the mix of patients seen in higher-level hospitals. In the next section we discuss the use of hospital resources to support primary health care. Finally, in the last section we discuss alternative institutional designs or treatment settings to reduce the cost and increase the effec-

tiveness of health services usually provided on an inpatient basis in hospitals.

Changing Hospital Referral Patterns

Hospitals provide a wide variety of services, from basic care to highly specialized diagnosis and treatment, depending on the technological capacity of the specific hospital. There is considerable duplication of services provided in different levels of facilities, and in many countries tertiary hospitals with advanced technical capacity devote considerable effort to delivering basic secondary and even primary curative care. These hospitals, however, are designed to treat complicated cases, and because of their composition of skilled staff and medical equipment the cost of treating a patient is often higher than the cost would be for the same type of patient in lower-level or specialized alternative facilities. This misuse of the capacity of tertiary hospitals affects outpatient as well as inpatient services. The outpatient departments of large tertiary care hospitals often suffer from overcrowding; this situation is exacerbated by the presence of a large number of patients who require only basic curative care that could be provided in a lower-cost setting such as a district hospital or health center. If patients requiring basic care could be shifted to medically appropriate lower-level or alternative treatment facilities, health system costs would be reduced and economic efficiency improved.

Analysis of discharge information by cause and severity of the illness or condition is needed to assess the potential for reallocation of the case load within the hospital system. In large tertiary facilities in urban centers, perhaps 75 percent of outpatient care and 30 to 50 percent of inpatient care could be effectively delivered in district hospitals or at lower levels. The average cost data presented in table 3-1 suggest the cost savings that could be obtained from improved referral. A comparison of the cost per outpatient visit or inpatient day between varying levels of hospital suggests that the savings could be on the order of 50 percent of current costs per patient shifted in a typical hospital system. Secondary (provincial-level) and tertiary hospitals absorb about 30 percent of government health expenditures (see chapter 2). If just 33 percent of patients could be shifted from upper- to lower-level hospitals, the total savings would be 5 percent of the current government health expenditures and 10 percent of government hospital expenditures. This example is only an approximation, because the average costs may differ from the applicable marginal costs required to shift users from upper- to lower-level hospitals. But the estimates, even if taken as rough, indicate that the order of magnitude of the savings could be considerable.

The potential for these savings is clearly recognized in many government and donor surveys of health sector needs (see, for example, Jolly and King 1966 and World Bank projects in Botswana (1984 Family Health Project) and Brazil (São Paulo Basic Health Care Project). Recognition of the need to improve the referral system is much easier, however, than implementation of the change has proved to be. A WHO Expert Committee on the Role of Hospitals at the First Referral Level cited many problems impeding the development of better referral systems (WHO 1987a; Paine and Tjam 1988), which can be summarized as follows:

- Overloading of hospitals with self-referrals or poorly judged referrals
- Barriers of distance or patient financial status that prevent the patient from carrying through with referral
- Lack of confidence in lower-level facilities because of perceived low quality of care
- Inadequate flow of information between the hospital and health center or health post levels
- Lack of organizational and management links between levels of care.

These problems are interrelated. For example, the overuse of secondary and tertiary facilities with inappropriate referrals is affected by the poor quality of lower-level services and the lack of organizational links and poor information. As a second example, the inadequate flow of information between service levels is related to a poorly integrated organizational framework, as well as the reverse. Overcoming these problems requires a set of interrelated policies. We briefly discuss three key policy areas—structure of the referral system, management coordination, and quality improvement—that address the problems identified above.

Structure of the Referral System

To function effectively in encouraging patients to use the level and type of service that fits their needs most efficiently, the referral system needs a carefully developed structure. This structure should be integrated with primary health care at all levels, and in the next section of the chapter we explicitly examine the role that hospitals can have in supporting primary health care. At each level of the system, the mix of patients to be seen and available services should be consistent and well defined. Referral criteria should be set out for major diagnostic characteristics and disease categories. Manuals, information, training, facility capacity, and management should support the referral policy.

Although these recommendations appear clear and reasonable, careful, detailed planning and analysis are needed to implement them. A working system requires a specific design that details referral policy and conforms to local circumstances, such as the pattern of diseases, population density, transportation and communication infrastructure, education and income of clients, facility capacity, and training of providers. The system should be designed in consultation with providers and consumers at all levels, from the community through professionals at the tertiary level.

Development of a formal referral policy. A useful starting point for the development of a formal referral policy is the supply and demand paradigm. This paradigm clearly recognizes that client demand is formed by medical needs, a perception of facility abilities to fill the needs, and perceived costs of services, and that supplied services and referral rules interact with this demand to determine the use of services at each level. Development of the policy should begin with an epidemiological analysis of the population to determine the relative prevalence and incidence of various diseases and conditions. Next, alternative methods of addressing these diseases and conditions should be defined and costed. At this stage of health sector planning, highly detailed cost studies are probably not required, but cost analysis should account for the scarcity of inputs such as physicians, pharmaceuticals, or highly trained medical technicians. The appropriate mixes of providers and of services to be available at different levels of the provider pyramid are those that generate the greatest population coverage at least cost. Population density is a key determinant of the appropriate composition of the provider network. A survey of existing referral patterns and facility use should be an integral part of the analysis and should include an accounting of the disease classification of patients treated; type of service received; distance of treatment center from household location; and referral history with regard to prevention, diagnostics, treatment, and outcome. This information, together with a review of existing procedures, staff training, manuals, and facility quality, can be used to inform the development of an improved system.

So far, the discussion has focused largely on health planning decisions. It should be recognized, however, that patients exercise considerable sovereignty in deciding where they will seek treatment. Mwabu (1989) notes that under certain conditions, a hierarchical referral system (such as that depicted by the pyramid) minimizes consumer costs of treating or preventing illnesses, and so cost-minimizing behavior on the part of consumers and health planners would lead them both to this common organization of health services. But the accuracy of the referral system as a model of consumer behavior is affected by the extent to which six

stringent assumptions, in addition to cost minimization, hold. The first assumption is that the same conditions can be treated at more than one level of the provider network. The second is that treatment costs for everyone are greater in higher-level facilities. The third, that the quality of services at all levels would be acceptable to patients. The fourth, that patients are well informed about the types of services available at each level of the network. The fifth, that patients could not bypass one level of the system without the consent of personnel at that (or lower) levels. The final assumption is that public providers would be the only source of medical care available.

The first assumption clearly holds for some conditions but not for others. Under conditions in which it does not, patients will not necessarily visit the cheapest treatment source. If the second, third, fourth, or fifth assumptions do not hold, patients will bypass the nearest facilities. Evidence on costs in different levels of hospitals suggests that the second assumption is generally an accurate depiction of the world. Inadequate consumer information and the perception of poor quality at lower-level facilities, however, often cause referral systems to function poorly. The enforcement of referral rules (the fifth assumption) varies across countries. The last assumption would ensure that follow-up care is provided in the public referral system. Mwabu's evidence from Kenya suggests that if nongovernmental alternatives are available, patients often leave the public provider network.

Expert assistance is needed to design a referral system. The skills needed are those of health planners, epidemiologists, medical doctors and nurses, economists, and sociologists. A nation's health authorities should convene a panel of clinicians and others to define the services to be provided at each level of the system. After the services are defined, it is useful to form a task force that includes representatives of consumers and producers at all levels to review the design. The referral system is more likely to be effective and followed by consumers and producers if it reflects their insight. In any case, whether through a formal process or not, draft designs for referral should be thoroughly discussed at all levels. Referral systems should also be tested in prototype, and mechanisms should be provided for review and adjustment before the system is extended to a large area.

At the hospital level, improving the referral system implies defining, on the basis of epidemiological and cost analyses, the specific services to be provided and, perhaps more important, not provided at the different levels of hospitals—for example, district; provincial, or secondary; and central, or tertiary. An example of this definition of services is given in table 6-1, in which the types of activities to be performed (and some not to be performed) are defined for three categories of services (diagnostic X ray, obstetrics, and surgery) at each type of hospital in a hypothetical

Table 6-1. Examples of Selected Services for Hospitals at Different Levels of a National Referral Network in a Low-Income Country

Level of hospital	*Diagnostic X ray*	*Obstetrics/maternity*	*Surgery*
I	High-cost procedures requiring specially trained personnel	Moderate- to high-risk pregnancies	Specialized and major operations with intensive care unit available
II	500-ma equipment and tests using contrast media, not CAT scans or MRI	Moderate- to high-risk pregnancies	Intermediate-level operations involving some degree of specialization
III	Routine exams using BRS technology, not contrast exams	Ordinary deliveries (which may also use lower-level facilities), outpatient pre- or postnatal care (coordinated with health centers); capacity to handle some birth difficulties, possibly cesarean sections	Basic surgery not requiring surgical specialists

Source: Authors.

country. The actual definition of the services to be supplied at different levels of hospitals in the referral network must, of course, be a country-specific decision. The theme of the table is that the more technically complex interventions and those that require staff with scarce, specialized skills and training should be performed at higher-level hospitals. These interventions are typically for conditions that appear relatively infrequently in the population. Although it is suggested that higher-level hospitals limit themselves (to the extent possible) to performing tertiary interventions, it is often not possible for them to do so because alternative lower-level facilities may not exist in large urban areas. In addition, it may be cost-effective for a tertiary hospital to have certain lower-level technologies at its disposal, such as BRS X-ray equipment.

Tools that can support the operation of the referral system have been identified by WHO (1987a). The referral system needs to be supported with manuals and training that specify the functions of different elements of the system and the relationships between them. For example, a referral manual (or a training course for administrators) should emphasize that a hospital should endeavor to limit its services to the skilled and technical care designated in order to fulfill its role in the system. The

hospital should discourage potential clients from consuming services that can be delivered more cost-effectively by community health centers or other nonhospital facilities. One way to do this is to reduce the extent to which such services are subsidized in hospitals. Communication between different levels of providers should ensure that patients who are properly referred to hospitals from lower levels are not subject to additional delays or duplicate investigations. Patients should be referred back to lower levels with full information when upper-level hospital services are no longer needed. Finally, the organization and management of the health system should reinforce the functioning of the referral system.

These tools and principles can be applied to alter referral patterns and improve the mix of services provided in hospitals. Lower-level hospitals should concentrate on providing a reliable supply of inpatient, ambulatory, and diagnostic services. The emphasis should be on basic services; specialty care should be left to tertiary centers. Particular attention should be focused on the specific diagnostic and treatment services that support primary health care. In addition to supporting primary care, basic hospitals have important relationships with tertiary hospitals. Basic hospitals should have the capacity to identify patients who require specialized care and the means to transfer these patients to appropriate facilities. They should also be able to care for patients who have received tertiary care and require recovery and rehabilitative services, either as inpatients or in follow-up ambulatory visits. The referral role of basic hospitals should thus be bidirectional.

Changing the mix of patients served in tertiary hospitals. As indicated above, a problem often faced by hospitals designed to provide highly specialized services is that relatively few of their patients require such services. There are three principal causes of this situation. First, the hospital serves the role of district (that is, first-level referral) hospital for a relatively large urban population. Second, oftentimes, there is no basic facility nearby that can lessen the load on this hospital. Third, the population is dissatisfied with lower-level facilities and believes that effective, quality care is available only at the tertiary hospital.

The first two causes are related and reflect a reality that comes from the urban location of the tertiary hospital. Adding a basic hospital nearby may not be feasible because of the additional recurrent costs it would entail, but the establishment of ambulatory care centers may be a possible option. Still, it is to be expected that a large proportion of patients in tertiary care hospitals are urban people in need of basic care, and perhaps little can be done to change the case mix of this group. The third cause is more amenable to policy. By improving the quality of lower-level hospitals and other treatment settings, the population will be given

credible alternatives to tertiary hospitals. Given these alternatives, tertiary hospitals can actively discourage self-referrals from outlying areas by enforcing policies on nonreferred admission or charging a penalty fee. Furthermore, patients who receive specialty care but require only basic rehabilitation and follow-up can be transferred back to a lower-level hospital nearer the patient's home.

If the tertiary hospital successfully alters its case mix by reducing the proportion of cases it treats that are not tertiary in nature, the likely result will be an increase in the cost per case at the hospital, because the resulting clinical composition of cases will increase in severity. The effect on unit costs at the lower-level hospitals to which these cases are transferred is uncertain. Average cost per case may increase or decrease, depending on the complexity of the additional cases and the extent of excess capacity previously existing in these hospitals. In this situation, the increase in "unit costs" at the tertiary hospital would reflect improved economic efficiency because a greater proportion of patients would be treated at less expensive, lower-level facilities. Facilities and their associated inputs would be more appropriately matched to patient needs. The overall economic efficiency of the health services would be improved, not just that of the tertiary hospital. Therefore, when a country evaluates the outcome of improving referral, efficiency should be assessed by consideration of the whole regional network rather than just one hospital budget.

Hospital financing and the referral system. Adequate levels of capital and recurrent resources are required for district hospitals and health centers to fulfill their designated roles in the referral system. The defined role will not be effective in practice if recurrent resource needs are continually underfunded. This is particularly true if the underfunding is greatest in basic facilities and results in shortages of supplies, insufficient staffing, and equipment breakdowns. When these conditions occur, consumers are more likely to refer themselves directly to tertiary centers in the belief that they will have a greater probability of receiving their desired treatment.

The financing system should support the referral design by creating appropriate provider and patient incentives. If the quality of services available in lower-level facilities is adequate, the fee structure can impose penalties on patients who refer themselves to a hospital when their condition actually warrants entry at a lower level in the system. Poor patients who use the referral system can be identified at lower levels (where such identification is easier), and if referral to more complex facilities is warranted, they can be exempted from hospital charges (Mwabu 1989). An estimate of the elasticity of demand for entry at different levels can provide information to determine a price that will

allow optimal cross-subsidization of patients using lower levels of services (see the appendix to chapter 4). A health insurer can also require referral compliance by withholding reimbursement or providing only partial reimbursement for patients who are not properly referred. For example, beneficiaries of the insurance program in Zaire's Bwamanda health zone (discussed in chapter 5) are responsible for 100 percent of hospital charges if they do not have a referral slip from a satellite health center. Finally, in prepaid capitation or fixed global budget systems, the provider will have an incentive to reinforce referral policy by using the most cost-effective level of services in managing patient care.

Some examples. The district hospital in Patan, Nepal, provides a good example of a facility that has committed itself to provision of basic services. It has 140 beds and serves a population of 210,000 in a region that does not have easy access to specialist hospitals. Nevertheless, planners resisted the urge to design and equip the hospital to provide specialist services and instead kept true to their prime objectives—provision of primary curative care, provision of secondary care through four departments (medicine, surgery, pediatrics, and obstetrics-gynecology), support for community health services, and training for health workers (Paine and Tjam 1988). The absence of specialists in lower-level hospitals may reflect a health system survival strategy for low-income countries that must cast a broad net of basic services and concentrate specialist services in a limited number of areas.

A second example of a hospital referral system that has acted to alter the role of lower-level hospitals to improve the efficiency of the mix of patients across several hospital levels is found in São Paulo, Brazil. In a World Bank project (São Paulo Basic Health Care) developed during the mid-1980s, several small hospitals in outlying barrios were created to provide credible basic inpatient services with close referral links to primary health care. None of the hospitals has an outpatient department. Instead, the hospital staff serves on rotation in satellite neighborhood health centers, designated the basic health unit, and screens for patients whose conditions warrant hospital care. The development of the new system was carried out in cooperation with INAMPS, the state social insurance system, to provide a higher quality of basic care and divert unwarranted use of expensive, higher-level services to be reimbursed by INAMPS.

Management Coordination and Organization

The organization of the health sector determines the relation between levels of hospitals and between hospitals and other parts of the health system. The organizational framework varies greatly across countries,

from those in which tertiary and large secondary hospitals are separately organized from lower-level facilities and most aspects of preventive and primary health care programs, to those in which regional organization and responsibilities emphasize the integration of different levels of facilities and services. Between these two extremes are regional or district administrative arrangements that emphasize coordination but fall short of complete management integration.

The arguments for and against integration or separation of hospitals parallel those raised in the larger debate on vertical as opposed to horizontal programs in health. There are strong arguments for both types of organization. Separation helps to maintain the financial autonomy of lower-level facilities and prevent larger facilities from absorbing recurrent revenues and staff ostensibly allocated to primary health care. Separation also provides some advantage in allowing management concentration on the reliable delivery of a more narrowly circumscribed set of activities. In many countries, however, these advantages may not offset the large benefits that integration provides in improving the functioning of referral and the effectiveness of primary health care delivery. An absence of organizational integration among hospitals at different levels impedes the functional coordination that is necessary to implement a referral policy that emphasizes the use of lower-level facilities to provide care.

One of the primary advantages of integration is that it internalizes incentives to contain cost and maintain the quality and flow of services at lower levels in the hospital and health services systems. Prevention programs should become important to managers responsible for the overall system because, if successful, they can reduce the use of expensive secondary and tertiary facilities. These managers also have an incentive to treat services at lower cost in district-level hospitals or in nonhospital facilities. An integrated system also adds incentives to improve logistics and maintain the flow of drugs and supplies to all levels of hospitals and other health facilities. Finally, the flow and use of information is improved by integration of planning and management across levels of hospitals.

Achievement of effective integration can be difficult, however, particularly if management skills are in scarce supply. The goal is to bring together the planning and budgeting process as well as broad aspects of management so that decisions affecting resource allocation at all levels can reflect the overall goal of providing the most effective services within the constraint of available resources. The diversity and complexity of hospitals presents an inherently exacting management problem, and the linking of hospital management across levels can add to the management burden. The difficulty is to provide sufficient links to create desirable incentives among different agencies in the supply of cost-effective

services and yet retain enough autonomy within the individual health unit to capture the efficiency of flexible decentralized management.

Coordination rather than integration may be a better management policy for the overall hospital and health system. The difference is more than semantic, because integration implies the formal merger of management functions under one administration. Coordination, in contrast, can be achieved without the imposition of a single, rigid, management structure. Innovative solutions to the management problem that emphasize coordination have been distilled by WHO (1987a) from the global experience. Among these solutions are the creation of area health boards to take responsibility for a range of hospital and nonhospital services, including, for example, secondary and district hospitals as well as lower-level services; the establishment of district health management teams that bring together health care professionals and managers from different levels and different functions within the system; the establishment of a community health department within hospitals; the location of the district health office within district hospitals; and the encouragement of community involvement in hospital managerial decisions.

Prior to the establishment of SUDS in Brazil (discussed in chapter 5), curative care services were financed and managed by INAMPS, the national insurance program, and preventive services were provided by the Ministry of Health. Under SUDS, the financing and management of curative and preventive services were unified and placed under the control of state and municipal authorities. By integrating management of all levels of personal health services, SUDS establishes an administrative framework that is consistent with an effective referral system. By unifying the functions and resources of INAMPS and the MOH at the state and local level, SUDS allows for a reallocation of health resources to a more cost-effective mix of preventive and curative care services and provides an administrative framework consistent with a well-coordinated referral network (World Bank 1988).

Quality

The quality and reliability of services provided in lower-level hospitals is critical to the overall functioning of the referral network. If patients have serious doubts that they will obtain the services they desire at a basic hospital, or if they do not trust that this level of hospital provides an adequate backup to their local primary care provider, they may bypass their local lower-cost providers and refer themselves directly to a tertiary hospital. Such activity adds to the economic inefficiency in the delivery system—tertiary facilities become overcrowded with basic cases, the fixed inputs of basic secondary hospitals are underused, and patients incur travel and waiting costs that would have been avoided if

local facilities provided a reliable supply of adequate quality services. An improvement in quality that shifts demand for basic services to basic facilities will result in a more economically efficient use of health sector resources and allow for greater coverage than was previously possible.

Quality and the market for health services. The perceived quality of health services acts together with price to determine an equilibrium quantity of those services demanded in the market. The critical question is whether the magnitude of the additional benefit resulting from improvements in quality that shift demand to lower-level facilities more than offsets the cost of the improvements. This is an empirical question that has recently begun to receive tentative answers. Policy simulations for Kenya (Mwabu and Mwangi 1986) and Nigeria (Akin and others 1991) that use empirically based parameters demonstrate that increases in fees that are used to improve the quality of services in lower-level facilities could actually lead to increased use. Mwabu and Mwangi estimate that the probability of government facility use will increase with the introduction of fees and an upgrade of quality to the level of nearby mission facilities. With an increase of fees from free to one Kenyan shilling per outpatient visit, the probability of use goes down by about 5 percent if there is no change in quality, but an upgrade in quality made possible by the fee increase would increase the probability of use by 25 percent. The basis for Mwabu and Mwangi's cost estimates is not made explicit. A similar finding, however, has been obtained for Ogun State in Nigeria using a facility demand function estimated by Akin and others (1991) and a cost function estimated by Wouters (forthcoming; discussed in chapter 3). A doubling of fees would mean less than a 4 percent decline in the probability of use, but if the fee increase were used to ensure the availability of essential drugs, the probability of use would increase by 35 percent.

This evidence of the willingness of users to pay for improved quality is expected to hold its general validity in many countries in which the price of services is currently low and poor quality is an impediment to appropriate use. The appendix at the end of this chapter provides an analytical framework that unites the demand and supply responses with quality and relates them to price changes. Most decisionmakers are aware of the role of prices in equilibrating service demand with available supply. It is important for managers and planners to be aware that reaction to quality can be as important as reaction to price in determining demand and equilibrating the market for health services.

The appendix identifies the role of quality as an equilibrating force in the market for health services and identifies the critical parameters to estimate the net effect of price and quality changes. The overall quantity response to a price increase, where the revenues derived are used to bolster quality, depends on the relative magnitudes of the price and

quality elasticities of demand and the cost elasticity of quality. If plausible values are used for these elasticities in a setting in which fees and quality are initially low and in which any increase in revenues is used to improve quality, it is estimated that a 10 percent increase in prices will have the net effect of increasing demand by 6 percent. Further research on the effect of quality on demand and cost is needed to establish the quantitative importance of these findings in a variety of settings. The findings in the appendix demonstrate, however, that the willingness to pay for quality is a robust conclusion over a wide range of plausible values for the underlying demand and supply elasticities.

The dimensions of quality improvements. There are several dimensions to the quality of hospital services, including availability of supplies, staff training, and type and condition of equipment, all of which affect the cost of services. The population's perception of an adequate quantity and reliable supply of pharmaceuticals and supplies is related to the demand for services as established in econometric studies of health services demand in Nigeria (Akin and others 1991) and elsewhere. An extensive and costly list of drugs is not required to bring about an increase in the use of lower-level facilities. The United Nations Children's Fund (UNICEF) program of essential drugs has demonstrated that basic health requirements can be met by a relatively limited list of pharmaceuticals, and WHO has developed suggested lists of essential pharmaceuticals for district hospitals and health centers. Patterns of demand for health facilities also illustrate that patients are sensitive to the training and availability of health facility staff. The occupancy rates of Indonesia's district hospitals are related to the available specialties. During the 1980s, the Indonesian government upgraded selected district hospitals to include staff with some added training in basic radiology, surgery, and obstetrics, and the demand in these hospitals grew significantly. Finally, equipment in lower-level hospitals need not be extensive to fill their basic referral role, but appropriate maintenance and the staff training to make proper use of it is required. The kind of equipment and pharmaceuticals that are selected for district hospitals will vary with the resource constraints and epidemiological needs in a given country. For countries of middle income or below, WHO has suggested norms for laboratory equipment, diagnostic imaging, and pharmaceuticals in a district hospital (WHO 1979; WHO 1987c; WHO 1986; and WHO 1985). These norms do not have an official WHO status but do suggest that it is possible to provide high-quality service, with regard to supplies and equipment, using relatively low-cost technologies.

Maintenance is a perennial problem at all levels in the hospital system in many countries but is especially critical at the lower levels, at which access to parts and routine maintenance are hampered by location and

communications. It was noted in a WHO meeting on maintenance of health equipment (WHO 1987b) that maintenance is a pervasive problem across geographic regions. In Brazil in the early 1980s, for example, an estimated 2 billion to 3 billion dollars of existing equipment was not working because of inadequate service, lack of parts, or simply a failure to install the equipment. In their study of medical equipment in Africa, Bloom and Temple-Bird (1988) note that preventive maintenance is the exception rather than the rule in many sub-Saharan countries. Both the WHO and the Bloom and Temple-Bird reports note that the neglect of maintenance is the result of insufficiently trained personnel, lack of funds for tools and parts, and poor organization of existing maintenance capacity with maintenance responsibilities spread over several agencies.

Mwabu (1989) suggests six reforms to improve the efficiency and equity of referral systems. Four of them relate specifically to improving the quality of services at the most basic units of the provider network. These four include increasing budget allocations to health centers and dispensaries; providing incentives for doctors based at hospitals to visit health centers regularly; strengthening the diagnostic capabilities of health centers; and ensuring that essential drugs are available in health centers and dispensaries. His other recommendations are to increase patient charges for self-referred use of hospitals and to abolish the general outpatient services provided in hospitals. This latter recommendation would be appropriate only if the reforms aimed at improving the quality and scope of nonhospital services were implemented, so that hospitals would need to provide only specialized services. Although eliminating the general outpatient function of hospitals would not be immediately feasible in most countries, the transfer of patients who do not require specialized services to less complex facilities located closer to their homes would result in greater efficiency in service use and is an appropriate long-term goal.

Using Hospital Resources to Support Primary Health Care

Improvements in the referral system are closely related to the use of hospitals to support primary health care. Although the bulk of primary health care activities are carried on outside the hospital through outreach programs, community service, education, and outpatient care, hospitals, nevertheless, are important contributors to the success of PHC programs. In particular, the role of district-level (first-level referral) hospitals in supporting PHC has become critical.

The full range of primary health care activities that can benefit from hospital involvement is long. A survey conducted in a sample of sixty-five hospitals in Delhi, India, and the surrounding region (Ghei 1985) suggests the extent of potential involvement. The survey, which is

summarized in table 6-2, revealed that hospitals were involved in administrative support, training of nurses, outpatient rehabilitation, and health education, as well as in the more obvious primary health care activities of health promotion, family welfare, and ambulatory care. The list of activities from the Delhi survey only suggests the number of different functions that hospitals might support but does not detail the form of support. The form of hospital support for primary health care can involve integrated (or coordinated) management, logistics, training, and, finally, credible referral.

Coordinated Regional Management

In the previous section on referral we discussed the importance of management coordination across referral levels. Here, we emphasize that the links should also be carefully made with PHC. The WHO Expert Committee on the Role of Hospitals at the First Referral Level (WHO 1987a) found that hospitals and PHC were administered separately in many countries, making coordination difficult. Even in some countries that make no formal distinction between hospitals and PHC, it has long been the practice to keep hospital management and information systems separate from those of community health services.

To bring about management coordination of primary health care, the WHO Expert Committee recommended that all local services, including hospitals at the first referral level, should relate themselves to a defined population.Within this population, communities should be involved in operating the "district primary health care complex," which includes district hospitals. It was further recommended that district hospitals, in consultation with their community and relevant agencies in the area, should have clearly defined referral, management, and support roles to play in their specific locality. Finally, these roles should be arranged to relate closely to other sectors of the economy, to higher health facility referral levels, and to various levels of government.

A serious impediment to carrying out the WHO recommendations is the separate administration and management of primary and secondary care found in many countries, which makes coordination of these levels of care a difficult process. Unified administration of all district health services, although not in itself sufficient to improve the referral system, would facilitate implementation of the strategy suggested by the WHO Expert Committee. Given a budget and a mission to provide health services to a district population, the administrators could face incentives somewhat similar to those of a prepaid capitated health plan, such as a health maintenance organization. The administrators of an HMO use the resources generated from premium payments to meet the health care demands of each participant in the plan. If these demands are met using

Table 6-2. Substantive Support for Selected Primary Health Care Functions in Sixty-five Delhi Hospitals
(percentage of hospitals involved with function)

Function	*Percentage of hospitals*
Health promotion	
Well-baby clinic	69
Education	62
Early detection programs	34
Marriage and motherhood	32
Curative health care	
Emergency services	46
Ambulatory care	100
Home care	17
Administrative support	
Planning and management	31
Office accommodation	37
General supplies (including pharmacy cold chain and the like)	37
Transport	32
Rehabilitation	
Inpatient	29
Outpatient	37
Home	8
Family welfare	
Education	60
Services	77
Supplies	75
Sterilization	69
Integrated hospital/community care	
Hospital responsibility for total health care in a designated geographical area	15
Hospital-based mobile clinics	17
Training	
Community health volunteers	2
Nurses	31
Paramedical staff	15
Management	5
Reorientation of professional staff	10
Preventive health care	
Drinking water and waste disposal	17
Communicable disease control	12
Immunization	69
Supplementary feeding	15

Source: Ghei 1985.

less than the HMO's premium income (less administrative expenses), the HMO makes a profit. If health service expenses exceed premium income, it operates at a loss. The analogy to an HMO is most appropriate if the district health service managers are held accountable for their resource allocation decisions, a condition that is not typical in publicly administered health systems. Management accountability is important because the effects of incentives are likely to be more powerful when managers have something to lose as well as something to gain as a result of the decisions they make.

Similarly, coordination of regional budgets and administration to include primary curative and preventive care as well as secondary and possibly even tertiary curative care would provide incentives to the administrators to organize the delivery of health services in a cost-effective manner. If the district is the unit of health management and operates under a fixed budget, management has an incentive to minimize expensive hospitalization through early detection and treatment of illnesses. The district would then be the level of administration at which attempts could be made to improve the allocation of resources across as well as within health facilities. To improve the referral system through support for PHC, it is not sufficient merely to establish the proper administrative structure, but it would seem to be necessary. As the principal health facility in a district, the hospital may be the logical place for the unified administration to be based. However, there may be some tendency for the hospital to be favored in the allocation of resources simply because of the presence of the administrators, who see the hospital's problems on a daily basis. Administrators should be conscious of this and guard against a hospital bias.

An example of a unified approach to the management of all health services for the population of a specific region is the Kasongo Public Health Project (de Béthune, Mercenier, and Van Balen 1985), which has been operating in eastern Zaire since 1971. The Kasongo area has one referral hospital plus a network of health centers. The health management team is based at the referral hospital and strives to cover its catchment population (about 250,000 people) with integrated and continuous care from an annual operating budget of about US$3 per person. Unified management of the area's health facilities has enabled the development of a coherent system that directs people toward the most appropriate setting and organizes service delivery using appropriate technologies.

The integration of the management of health services at Kasongo has resulted in many policies that have improved the functional relationship between the hospital and primary health care activities. The roles of the health centers and the hospital are clearly defined and relate to the severity of patients' conditions. Some complex diagnostic tasks and a

few appropriate surgical procedures have thus been entrusted entirely to nurses. Downward referrals to health centers from the hospital are routine, and home visits (when needed) are included in the package of services provided to the referred patient. The main pharmacy, which serves all the health facilities in the area, is located at the hospital, but priority in ordering and distribution is systematically given to the health centers. The financing system reinforces the primary health care structure. Referral and hospitalization charges are included in the initial fee paid by the patients on their first visit to a health center, whereas self-referred patients in the hospital's outpatient department face higher fees, long lines, and systematic encouragement to use a health center. Hospital length of stay has been reduced as more patients whose recovery can be managed at health centers are treated at that level of care.

Logistics

Regardless of the administrative structure of the district health services, a hospital can support primary health care activities in its region by assisting with the planning, budgeting, procurement, and distribution of needed materials to the PHC network. According to Milner (1985), the purpose of these logistical activities is to provide the necessary equipment and supplies, in correct amounts, to the correct health facility, health provider, or health program, and to do so in a way that minimizes costs. In many countries, hospitals are intermediate facilities in the procurement and distribution of items to PHC facilities and programs. Items are initially purchased and stored centrally, in bulk, and distributed to intermediate facilities (that is, hospitals) for their use and for redistribution to local facilities and programs. With limited stocks of supplies, however, the lion's share of these items typically goes to the intermediate facility.

The type of transport system best suited for support of primary health care depends on the means available to a country, the geographical accessibility of a region or district, and climatic factors. Because the viability of a transport system is somewhat dependent on policies outside the purview of the Ministry of Health, such as those determining road conditions or vehicle maintenance, intersectoral cooperation in the planning of public works projects is necessary to ensure that the needs of the health transport network are considered. Milner described how an effective transportation service can support the health system in a number of ways:

- provision of needed supplies and equipment—including movement of supplies from manufacturers or port of entry to the central stores, and distribution of supplies throughout all levels of the health service

- training—transport of supervisors and trainers from hospitals to satellite facilities, or bringing community health workers to hospitals for training sessions
- patient referral—bidirectional transport of patients across levels of providers in the referral network, including back to the community.

It is important for relatively isolated health posts to be able to communicate regularly and, on occasion, rapidly with their referral hospital. In rural areas without reliable telephone services, radio communications may be needed to support the education of community health workers by regular consultation with hospital staff and to inform the hospital of sudden changes in situations affecting health or the need for replenishment of specific supplies.

Training

The attitudes of those working in hospitals, influenced by their education and training, are often quite distant from what is needed for a community-based approach. Improvement of PHC systems involves retraining of hospital personnel to increase their awareness of the importance of PHC activities and to provide them with the supplemental skills necessary to carry out PHC tasks (Dean 1981). In addition, once trained, hospital staff can provide an appropriate base for regional training of outreach workers. The literature on the connection between hospitals and primary health care (Walker 1986) repeatedly emphasizes that relatively highly trained hospital staff members, such as physicians and registered nurses, can support primary health care activities in their region by providing continuing or in-service training to lesser-trained community health workers. This training can be done either in a hospital setting or as part of the outreach services provided by a hospital to its catchment community. Several studies (Kleczkowski, Elling, and Smith 1984; Carreon 1982; and Pust 1985) also make strong cases for the desirability of outreach supervisory and training visits.

The relationship between the Community Primary Health Care Program and the Patan Hospital in Lalitpur, Nepal, provides many opportunities for hospital staff to engage in educational activities. When patients are referred from the hospital back to health centers closer to their home, the hospital doctor sends written comments concerning the previous and future treatment for the patient, in addition to his or her diagnosis. Health assistants come to the hospital each month for lectures, and physicians from the hospital also travel to health posts. In this way, the hospital's doctors not only provide some on-site education to the lesser-trained health assistants, they also gain a better understanding of the conditions faced at the health posts (Gsellman 1985).

The Bethesda Tomohon Hospital in North Sulawesi, Indonesia, is involved in the training of volunteer health workers (VHWs), who provide basic primary health care to the villages from which they are drawn. The hospital staff members who provide the training to VHWs are themselves given prior training or retraining to impart knowledge that, until recently, has not been a part of medical training. Important components of the training program for primary health care integration in the Bethesda Tomohon Hospital include education concerning prevailing health problems and methods of prevention and control; nutrition education, available foodstuffs, and food preparation; basic sanitation and safe water supply; maternal and child health, including family planning; immunization against major infectious diseases; prevention and control of locally endemic diseases, particularly tuberculosis; appropriate treatment of common diseases and accidents, such as the use of oral rehydration therapy and basic first aid; and the use of essential drugs (Supit 1985).

The hospital can, thus, be both a base and a source of skilled personnel for continuing education and training of community health workers. The district will benefit by the training, advice, and supervision that highly qualified health professionals from the hospital can provide to PHC workers. In addition, the district health system can be the setting for university or other professional education programs to bring the importance of PHC directly into the training of future health professionals (WHO 1987a).

Diagnostic Support and the Credibility of the Referral System

The effectiveness of a referral system depends on the population's willingness to use lower-level facilities as a point of entry into the health system. As was emphasized in the first section, this willingness depends on individual perceptions of the quality of services likely to be received. Laboratory tests are an especially important service linked to quality that can be provided at the primary care level. Typically, the test is taken at the primary level, and the specimen is sent to a hospital laboratory for analysis. The results are then sent back to the primary care facility. If specimens are often lost or the results do not come back in a timely manner, people may bypass the health center and go directly to the hospital. This may be perceived by health planners as resulting in an inefficient pattern of health sector resource use, but it is perfectly rational behavior on the part of a consumer who has lost faith in the referral system. Providing accurate and timely diagnostic support to primary health care is an important means by which the credibility of the entire network of providers can be increased in the eyes of the public. This credibility is necessary for a referral system to function effectively.

The Ministry of Health in Jamaica is seeking to upgrade the laboratory and X-ray diagnostic support services throughout its network of hospitals, from district to tertiary facilities. One of the reasons for doing so is to improve the capacity of hospitals to support primary health care (GOJMOH 1987). Such improvements in the hospitals are expected to facilitate early diagnosis of certain conditions, support clinical strategies to improve quality and continuity of care, and strengthen linkages between the primary and secondary levels. The provision of reliable diagnostic services at the first referral level is an important means by which hospitals can increase the credibility of primary care in the eyes of the public and encourage a more economically efficient pattern of service and facility use.

Some Examples of Primary Health Care Activities Supported by Hospitals

Hospitals can play an important role in promoting all primary health care activities through the forms of involvement outlined above. For some activities, hospital services are involved directly through clinical referral support. In these activities, the health services provided by the hospital are directly integrated with the structure of the PHC program. Some important examples are provided by programs for breast-feeding, acute respiratory infection, safe motherhood, and cervical cancer screening.

Breast-feeding. Promotion of breast-feeding, which is a core component of the UNICEF and WHO child survival strategy, is often neglected in hospitals. An innovative program in the Philippines is demonstrating the importance of breast-feeding as a means of reducing hospital costs as well as improving child survival. In 1979, in response to a decline in breast-feeding in the Philippines, the Department of Health instituted a policy requiring rooming-in within maternity facilities and government hospitals (Wellstart 1990). In 1986 the Philippine milk code was adopted, restricting the sale of substitutes for breast milk directly to the public and through hospitals. As implemented in the hospital with the largest maternity service in the Philippines, the program involves inpatient breast-feeding with the child in the mother's room twenty-four hours a day, immediate nursing after delivery, and follow-up with special clinics that support outpatient lactation management. A detailed study of the program revealed that, in addition to the benefits of increased breast-feeding, it is saving a net amount equal to 8 percent of the hospital budget per year. The savings derived from reduced costs of formula, bottle-feeding sets and associated supplies, and savings in nursing staff time.

Acute respiratory infection. Acute lower respiratory infections (ARI) are a major cause of child mortality. Development of new, low-cost diagnostic and treatment alternatives for bacterial-based infections has created the potential for special programs targeted at acute respiratory infections. Succinctly stated, the programs include three elements: identification of infection, treatment at home or through outpatient care where possible, and referral of severe cases to higher-level facilities. District-level hospitals can play an important role in developing and sustaining an ARI program. The hospital plays an obvious role in referral but is also involved in the other elements. The need for timely and careful diagnosis of infection and adherence to the treatment regime places an emphasis on training and information for the mother and outreach workers. A regular and sufficient supply of antibiotics is also essential to the program. Potentially critical to its success is hospital involvement in training, logistics, and supervision. Finally, referral of severe cases to higher-level facilities requires adequate drugs, oxygen, and laboratory and radiological capacity in the district hospital.

The cost-effectiveness of ARI programs has not yet been carefully established, especially across urban and rural settings with varied infrastructure. Problems of identification of false positives, failure to comply with treatment instructions, diversion of program antibiotics to other uses, and a lack of adequate backup facilities may limit their impact in some communities. In other communities, with sufficiently motivated and trained staff, adequate supplies, and adequate supervision and management, the programs may prove highly cost-effective. Ultimately, the cost-effectiveness of ARI programs in specific environments may depend on the level of development of infrastructure, including hospitals.

Safe motherhood. Worldwide, an estimated half million women die in pregnancy and childbirth annually. Almost all these deaths occur in developing countries (Starrs 1987). A comparison of the less than 15 maternal deaths for every 100,000 live births occurring in industrial countries with the 450 per 100,000 occurring in developing countries makes it apparent that many of these maternal deaths could be prevented. An important initiative by international development agencies collaborating with individual ministries of health has been mounted to promote safe motherhood. The major elements of the initiative are promotion of a full range of family planning services; prenatal care, including assessment of maternal risk; safe, routine delivery for low-risk mothers; and hospital delivery for mothers with complications.

Hospitals, especially at the district level, play an indirect part in risk assessment (especially through training) and a direct part in the other three elements. To support family planning, hospitals need adequate

facilities for abortion and sterilization. In low-resource settings, hospital delivery of low-risk mothers is not cost-effective, but delivery by trained birth attendants in the home, in bedded health centers, or in birth centers outside of hospitals can substantially reduce maternal and child death caused by septic or traumatic delivery. A functioning transportation system is required to provide timely access to hospital services for pregnant women with complications. Deliveries requiring referral services vary from 5 to 15 percent of births. During pregnancy or delivery, women with disease or complications may require hospitals with a reliable blood bank and the capacity to perform safe cesarean sections. A small proportion of high-risk deliveries will need further referral to higher-level hospitals.

The cost-effectiveness of safe motherhood programs has not been assessed in an operational setting and with all the elements of the full program. Cost-effectiveness of prenatal screening and risk assessment followed by appropriate delivery is strongly suggested by the prototypical model set out by Herz and Measham (1987) and used for the cost-effectiveness comparisons in table 2-9. Given the potentially critical importance of hospitals in all elements of the program, however, the cost-effectiveness of safe motherhood programs in alternative settings needs to be further examined. In particular, the sensitivity of the cost-effectiveness to the availability of hospital services at the district level needs to be established.

Cervical cancer screening. Globally, approximately 500,000 new cases of cervical cancer occur per year, and 80 percent of these occur in developing countries.[1] In the cervix, progression from normal cells to cancer usually occurs slowly (during a period of fifteen to thirty years) and in phases of dysplasia, carcinoma in situ, and finally invasive cancer. The long period during which the condition can be detected at a preinvasive stage makes screening a practical and effective possibility. In the United States, Canada, and Western Europe, widespread use of the Pap test for screening has contributed to a fall in mortality from cervical cancer.

In developing countries the success of a cervical cancer screening program depends on the level of development of supporting infrastructure to provide accurate testing and timely and effective follow-up of confirmed cancer. A Pap smear can be taken at a health post, but access to laboratory facilities and communication back to the health post and then to the patient are potential problems in environments that have a poorly developed infrastructure. District hospitals can play an important role in providing adequate laboratory analysis of the smear and, especially if they are well integrated into the larger primary health care system, can also facilitate communication and supervision required for follow-up. When carried out efficiently, screening for cervical cancer by

use of the Pap smear allows early diagnosis of the disease and therefore the use, in a health center or first-level referral hospital, of a low-cost and low-risk procedure (cone biopsy) to remove cells and prevent occurrence or spread of invasive cancer. Cervical cancer diagnosed at later stages requires more costly surgical intervention in higher-level facilities.

As summarized in chapter 2 (see table 2-9), cervical cancer screening is a cost-effective intervention in middle-income countries. It is also potentially cost-effective in low-income countries provided that adequate supporting hospital infrastructure exists. As with the other PHC activities noted in this section, clarification of the cost-effectiveness of cervical cancer screening in alternative environments, especially with regard to hospitals, is necessary. Unlike the other activities cited, however, the existing literature on cost-effectiveness of cervical cancer screening provides a good basis for further studies (see especially, Parkin and Moss [1986] and Eddy [1986]).

Alternatives to Inpatient Hospital Care

The need for low-cost alternatives to the existing use of hospital care is motivated, in part, by the economic implications of demographic and epidemiological projections that demonstrate that the potential demand for hospital services will continue to outstrip available resources for the foreseeable future. (See, for example, projections for China [Bumgarner 1992] and Brazil [Briscoe 1990].) The development of innovative, lower-cost alternatives to inpatient hospital care is consistent with the logic underlying the referral pyramid. If a specific technology or treatment setting is relatively low cost and the disease to which the alternative applies occurs with sufficient frequency in the population, it will be a cost-effective alternative, by the same rationale that health centers are more cost-effective than hospitals for treatment of certain very basic conditions. The motivation is not only economic, however; alternative modes of delivery may also be more effective, accessible, or humane. We consider several examples below of alternatives that potentially cost less and are more effective than hospital care. The alternatives fall into two basic categories. First are alternatives that involve changing treatment protocols or use treatments that substitute new and appropriate technology for inpatient care; and second are alternatives that replace care in tertiary facilities with lower-level treatment and palliation.

Treatment Alternatives

Considerable cost savings are possible from revising treatment to exploit new technologies or procedures, and many newer and cheaper procedures result from improved pharmaceuticals or equipment. But reexam-

ination of older treatment protocols has also brought about a revision in accepted medical views on the use of home care and ambulatory services in place of extended inpatient care, and many new treatment procedures do not require expensive new equipment. Medical personnel trained many years ago and not in close communication with medical change may not have incorporated recent cost-effective techniques in their everyday practice. Also, many of the improved treatment alternatives are not easily changed by an individual hospital or practitioner but require functional changes across health care institutions and, thus, must be introduced as part of the health planning process.

Planning and implementing new treatment procedures is detailed and technical, and no general rules for change can be given here. Extensive experience does exist worldwide with new protocols that hold out the promise of more efficient resource use in developing countries from the adoption and adaptation of existing technology. New technology is a two-edged sword that can either increase cost and reduce the equity and efficiency of the overall system or reduce cost and increase the accessibility of health care. Analysis of the cost-effectiveness of potential new technology is required to establish the desirability of change. In countries isolated from the mainstream of new medicine, it is recommended that expert assistance be used to survey existing practice and suggest new modes for improving the cost-effectiveness of treatment practices. Actual change in practice should be carried out cautiously with experimental implementation and evaluated before systemwide adoption.

As means to reduce the use of secondary and tertiary inpatient care, new treatment procedures are particularly promising in the following areas:

- *Low-risk deliveries.* As noted above, improved screening of pregnancies may allow low-risk deliveries to be shifted to urban maternity centers. The effectiveness of this strategy is related to the existence of a reliable emergency transport system so that cases with unforeseen complications can be brought to a nearby hospital.
- *Tropical and communicable disease treatment.* New pharmaceuticals make it possible to treat some stages of selected communicable diseases through outpatient care. Examples can be found in treatment for malaria, tuberculosis, and leprosy.
- *Reduction of unnecessary surgical procedures.* Cesarean sections, tonsillectomies, and hysterectomies are common examples of procedures that have been overused in the past. Overuse of these procedures is closely associated with the design of health care financing.
- *Ambulatory surgery for selected procedures.* Ambulatory surgery has become common in industrial countries. The practicality of ambulatory surgery in developing countries depends on the geographic

accessibility of health services and the quality of the home recovery environment. Among the many procedures that have become more common in ambulatory care are inguinal and umbilical hernia; dermatological surgery; breast biopsy; surgery for cervical neoplasia; many ear, nose, and throat procedures; vasectomies and tubal ligation; and orthopedics.

- *Outpatient treatment of chronic disease.* Heart disease, diabetes, and cancer provide examples of diseases for which outpatient protocols and home care have increasingly replaced hospital inpatient care.

These possibilities suggest that innovative treatments that can reduce the cost and use of hospital services cover a wide range of activities. Outpatient hospitals, tuberculosis treatment, and oral rehydration therapy provide especially clear and contrasting examples of cost-effective alternative treatment methodologies.

Outpatient hospitals. Greater use of outpatient services to treat and diagnose disease that has been more traditionally addressed through inpatient services can provide cost savings through reduction of overhead and staffing. New approaches to surgery provide an example of the cost savings that are possible through substitution of outpatient for inpatient care. In Europe and the United States, many basic surgical procedures, which formerly required short stays in the hospital for the patients, are provided in specialized facilities called ambulatory surgery centers. In these centers, the surgical procedure is performed, and if there are no complications, the patient is released the same day. Surgery is mentioned because it provides a striking instance of how the mode of service delivery can be restructured to achieve greater efficiency. A range of other services, however, are also being made available in outpatient facilities in OECD countries. Similar facilities have been made available in some parts of the developing world.

Outpatient hospitals in Cali, Colombia, are examples (Vélez-Gil and Pardo de Vélez 1985). These hospitals provide a wide range of services: normal maternity and delivery care, first-contact emergency services, diagnostic services with full X-ray and laboratory facilities, ambulatory surgery, and outpatient treatment and follow-up for a broad spectrum of conditions. The importance of these centers lies in their capacity to provide a wide range of services in a more cost-effective location than the inpatient hospital yet retain credibility and quality from the point of view of the patient.

As an indicator of the relative cost-effectiveness of the outpatient hospitals in Cali, Shepard and others (1991) evaluated the cost-effectiveness of the ambulatory surgery centers that are an integral part of the outpatient hospital design. These ambulatory surgery centers were cre-

ated in response to findings that nearly 70 percent of surgeries performed at the tertiary hospital in the area did not require highly specialized services. Analysis demonstrated that ambulatory surgery centers located in nearby urban or peri-urban areas could provide simple elective procedures at far lower cost. For low-risk patients, inpatient procedures and follow-up were replaced by ambulatory surgery with postoperative home care. Shepard and his colleagues found that the average cost per inguinal herniorrhaphy patient in the ambulatory surgery centers was about one-quarter that of those treated as hospital inpatients. Furthermore, the quality of care (based on an assessment of postoperative results and complication rates) provided in hospitals was not found to be better than that provided in the centers. There may have been differences in case mix between the sets of patients sampled at each facility, however, because the mean age of the hospital patients was twelve years greater than that of the patients in the surgery centers. Nevertheless, the magnitude of the cost difference clearly points to the conclusion that the cost (adjusted for quality and case mix) per inguinal herniorrhaphy patient in the ambulatory surgery center is considerably less than in the hospital.

Several factors contribute to the success of the center in Cali, Colombia, that may cause difficulties in replicating its benefits elsewhere. Colombian health facilities have an excellent referral system, which works especially well in Cali because of its urban location and its relatively high standard of living. Furthermore, Colombia is reported to have an "oversupply" of physicians. It would probably be difficult to sustain ambulatory surgery centers in a rural area of a poor country suffering from a scarcity of physicians. Additionally, adverse home conditions that could impede recovery can be an important limitation to the wide use of ambulatory surgery in the least industrialized countries. In these circumstances, having separate ambulatory surgery centers in addition to hospitals would be less likely to be as cost-effective for a health ministry.

It is uncertain how the savings resulting from the use of ambulatory surgery centers in Colombia are being used. However, one benefit that seems apparent is that the operating theater of the nearby tertiary hospital can now handle a greater number of more complicated cases than was previously possible (though this was not explicitly assessed as part of the study). Studies similar to this one, performed on a prospective basis or as evaluations of pilot schemes, should be conducted by countries as part of their health sector investment planning activities to determine the usefulness of outpatient hospitals, and the range of services to be made available, in specific contexts.

Tuberculosis treatment. Until recently, the conventional treatment for tuberculosis was long-term therapy that involved approximately three months of inpatient care followed by about fifteen months of outpatient

care. Treatment was based on the antimycobacterial drugs thiacetazone and isoniazid. In the 1970s the introduction of new antimycobacterial drugs, ethambutol and rifampicin, made shorter-term treatments of six to eight months possible. The cost-effectiveness of rifampicin and ethambutol is not self-evident, given that isoniazid is only a fraction of the cost of the short-treatment drugs, and the older therapy requiring long-term treatment remains in use in regions of many industrial countries. The high cost of rifampicin and ethambutol is more than offset in most countries, however, by the higher compliance of the patient in the short-term treatments and the substitution of ambulatory for inpatient care.

The unit cost of tuberculosis treatment is high, especially if the long-term inpatient protocol is used, and in countries in which it is endemic the cost of treatment of this one disease can amount to more than 10 percent of the total public expenditure on hospitals. Analysis of the cost-effectiveness of tuberculosis treatment in Botswana (Barnum 1986) revealed that after improved compliance and a shift to ambulatory care were taken into account, the cost of the short-term treatment is less than one-third to one-half the cost of isoniazid-based regime per person effectively treated. Any given program is necessarily a mix of treatment strategies, depending on the accessibility of services, health care practices and customs affecting compliance, and the prevalence of resistant strains. Data for Botswana for 1982 suggest that the adoption of the short-term ambulatory treatment for 80 percent of patients would reduce the total health expenditure for tuberculosis by two-thirds, and the number of people that complied with the treatment and were cured would double.

Oral rehydration therapy. Diarrhea is a major cause of early childhood death throughout the developing world. Death occurs primarily as the result of dehydration. Until the early seventies, the standard treatment was to use intravenous (IV) rehydration therapy in an institutional setting. The high cost of IV therapy and a lack of timely access to a hospital or health center limited its usefulness in combating diarrheal mortality. More recently the introduction of oral rehydration therapy (ORT), which consists of a simple solution of sodium and potassium salts in water administered in the home or other outpatient setting, has dramatically increased the effectiveness of treatment and reduced the need for the higher-cost intravenous therapy.

Many studies have confirmed the very high cost-effectiveness of ORT. Horton and Claquin (1982) used 1980 data for the diarrheal program in Matlab, Bangladesh, to compare the cost of diarrheal treatment in a large health center with that in a health post. In both cases the therapy involved a combination of IV and ORT as required in each case, but with

the primary difference that the larger facility relied more heavily on IV therapy. They found that the cost per death prevented in the health post was less than one-fifth that of the large health center. Still greater savings are possible through the use of home-based ORT supported by outreach workers or outpatient clinics. A study of ORT therapy delivered through alternative modes in Egypt (Shepard, Brenzel, and Nemeth 1986, based on a reinterpretation of data analyzed by Mobarak and others 1981) provides a comparison of ORT used in the home with hospital clinic-based therapy. The cost per death averted by using premixed salts in home therapy was less than one-tenth of the cost of care in a control group in which the primary therapy was hospital-based IV treatment. Thus, considerable benefits in saved resources and increased effectiveness can be obtained from shifting treatment from a large to a small facility, and still greater gains can be realized from the substitution of home-based ORT therapy for clinic-based IV treatment.

The example of the treatment for diarrhea illustrates two important points. First, it provides another illustration of the close link between hospital services and the design of primary health care programs. Oral rehydration therapy is a core intervention in primary health care efforts to improve child health. Second, it demonstrates that technological change supporting more efficient use of hospital resources can be simple and relatively inexpensive.

Lower-Level Treatment and Palliation for Chronic Diseases

Alternatives to the use of hospitals designed to provide acute care services are needed to provide more cost-effective and more humane palliative and rehabilitative services to those suffering from chronic conditions. Although most of these conditions, such as cancer, are non-communicable, the rise of the AIDS epidemic is also placing tremendous burdens on the existing health facility infrastructure. Lower-cost settings and the conversion of hospitals from acute to chronic care facilities are alternatives that have been used in some developing countries.

Use of district hospitals as extended care facilities. In some areas, district hospitals whose capacity is underused have the potential of being used for extended care. Basic secondary facilities that serve as extended care institutions and provide recovery and rehabilitative care to patients suffering from chronic conditions could be an important complement to a coordinated referral network among different levels of hospitals. Also, underused secondary care facilities could be adapted to provide hospice in addition to chronic care. Such institutions could remove from tertiary care those patients for whom the higher-cost technical capacity is not required or effective. As the demographic and epidemiological transi-

tion that has started in many countries continues, a greater proportion of patients are likely to suffer from chronic diseases, and extended care and rehabilitation will therefore become increasingly important. Currently, in many countries, acute care hospitals are being inappropriately used to provide both acute and chronic care because of the lack of facilities that can provide rehabilitation services (Young 1989). A referral system that provides early identification of those who need long-term care and directs them to an appropriate lower-level facility may reduce the burden on larger, short-term hospitals. Reducing the rehabilitation function of tertiary hospitals should allow for some reallocation of staff and equipment across facilities, which may result in improved sectoral productivity.

One strategy that has been implemented in China is for groups of hospitals to coordinate their services used in the treatment of chronic care in an attempt to reduce duplication and waste of resources. Chronically ill patients are transferred from tertiary-level to district hospitals for recovery and rehabilitative services (Young 1989). Such arrangements hold out the possibility of a more economically efficient allocation of resources within a health system, but detailed quantitative evaluations need to be conducted to establish cost-effectiveness.

Alternative institutional settings. The technology exists to provide much of the treatment and care needed by persons with chronic diseases outside of a hospital. Also, for the terminally ill, hospice care provided in either a home or institution may be a humane alternative to hospitalization. Home care and the use of neighborhood or community centers provide the advantages of retaining the proximity of the patient to family and community and potentially lowering the cost of care. However, although experimental programs exist in various countries, no clear analysis of the cost-effectiveness of these alternatives has been conducted. For example, in a few areas China has experimented with the use of home beds as well as community centers for those suffering from chronic diseases, but no cost-effectiveness studies have been done to establish the cost advantages of these arrangements over the use of hospitals. Nevertheless, it is likely, particularly for countries that have overcrowded hospitals, that the use of alternative settings would relieve some of the burden from hospitals designed to deliver acute care and would result in a more economically efficient use of a country's health sector resources.

Care provided at home by family members may be a practical alternative for delivering some rehabilitative services such as chest percussion and drainage for emphysema patients. Analysis of the cost-effectiveness of this alternative in a particular setting is needed to establish the usefulness of a home care policy. Family members giving care may not

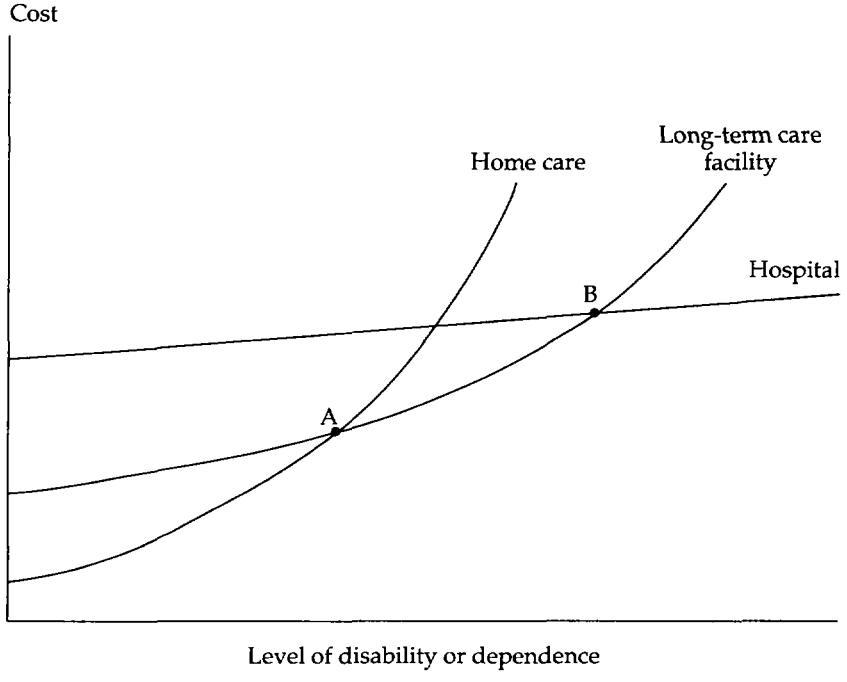
be available cheaply if they must forgo productive labor to tend to their sick relative. Another home care alternative is to have a trained health worker (for example, a nurse) visit the homes of persons requiring rehabilitative care or monitoring. Again, this may be practical in some countries, but the viability of such a program depends on the availability of trained personnel and the population density of each region that is considering such a visiting service.

Another alternative to hospital-based management of patients with chronic conditions is care provided to a group of disabled persons in a neighborhood or community center. It would seem to be economical to have a single care provider, such as a physical or respiratory therapist, deliver services to several patients with similar symptoms or disabilities. A community facility could potentially serve as an adult day care center, providing health assessment, monitoring, nutrition counseling, personal care, and various forms of rehabilitative therapies. Possible problems with such centers involve the financing of their recurrent costs (over and above the costs of health personnel) and transport of disabled persons to the centers on a regular basis. In addition, community-based rehabilitation programs may not have large or powerful enough constituencies to provide the support needed for the development of such programs (Young 1989).

Drummond (1980) has suggested that the cost of caring for the dependent elderly as a function of disability will differ markedly by institutional setting (figure 6-2). Although Drummond's intended context is care for the elderly in an industrial country, the conception of cost as a function of disability is applicable more generally to care for the chronically disabled in developing countries. As depicted in the diagram, the cost of home care is initially low but increases more rapidly with disability than does the cost of care in an institution in which the equipment and training are designed to cope with disability. The critical points to be discovered through cost-effectiveness research are the point at which home care becomes less cost-effective than community residential care (point A), and the point at which disability is so severe that hospital (or possibly, hospice) care becomes most cost-effective (point B). The diagram illustrates the complexity of answering the question of relative cost-effectiveness of care in alternative settings and suggests that the best solution will be a mix of care possibilities. Development of appropriate disability scales and collection of cost data in alternative settings is needed to identify the likely points of intersection.

Acquired immunodeficiency syndrome. Consideration of alternative settings of care has been made more urgent by the burden on health facilities that, in some areas, has already resulted from the epidemic of acquired immunodeficiency syndrome (AIDS). A Panos Institute dossier

Figure 6-2. Variation in Cost with Level of Disability and Form of Care for Chronic Disease



Source: Drummond 1980.

(Sabatier and Tinker 1989) reported that one-quarter to one-third of beds in some central African hospitals were occupied by AIDS patients. The burden will certainly increase, as persons infected with the human immunodeficiency virus (HIV) develop AIDS. According to WHO estimates, the annual number of new AIDS cases in sub-Saharan Africa is projected to grow from less than 300,000 in 1991 to nearly 450,000 in 1994. The total number of cases in Asia is projected to increase more than tenfold during the same period (WHO/GPA 1992).

Episodes of illness in AIDS patients often result in repeated hospitalizations. Yet the wide range of estimated treatment costs per AIDS case exhibited in both Tanzania and Zaire suggests that low-cost (presumably community and home care) and high-cost (hospital care) settings exist for the care of such patients (Over and others 1988). There is not much available information on alternative treatment settings for AIDS patients in developing countries, however. A report summarizing the medical,

public health, social science, and popular literature on AIDS in Africa (Johnson and Pond 1988, p. 5) noted the need for such information:

> With so many preventable and 100% curable health problems in Africa it might be argued that it is a mistake to divert scarce resources to management of a 100% fatal disease. . . . It is inevitable that medical treatment of AIDS will consume a growing part of African health care budgets. Much more needs to be written about humane yet cost-saving approaches to managing these patients both in hospitals and in the community.

Sewankambo (1989) advocates a strategy to deal with HIV-induced disease similar to those strategies used for chronic diseases such as diabetes mellitus and hypertension, wherein the patients take an active role in their own treatment. Three levels of coordinated care are suggested. First and foremost, home and community-based care can give patients cost-effective treatment while enabling them to lead as normal a life as is possible. Most care should be provided by members of the family of the infected person, and family caregivers must be educated about HIV transmission and prevention. Second, specialized HIV clinics should be established in areas in which infection rates justify their existence. Finally, inpatient care should be a last resort and used only if outpatient treatment is not feasible and if the patient would benefit from hospitalization. Although Sewankambo stresses the need to prevent nosocomial HIV transmission through infection control procedures, he rejects the idea of establishing separate AIDS wards because they would be likely to stigmatize patients and increase their sense of isolation.

Palliation. An important part of chronic disease care, especially for the elderly, is concerned with palliation. The pain of cancer or discomfort of chronic obstructive pulmonary disease leads patients to seek hospital care even though such care may not offer any prospect of cure. An alternative that may be at once less costly and more humane is to provide palliative outpatient support. Cancer provides an example. The World Health Organization recommends a three-stage analgesic therapy that moves progressively from nonopioids for mild pain, through weak opioids to strong opioids for moderate to intense pain (Swerdlow and Stjernswärd 1982). The cost of palliation, including drugs and staff time, for an average duration of therapy of ninety days represents 3 to 7 percent of the cost of alternative tertiary hospital treatment of advanced-stage cancer (Barnum and Greenberg 1991).

The failure to use palliative alternatives for chronic disease results from a lack of recognition by health care professionals that effective and humane alternatives exist to hospital care. A lack of availability of the required drugs, unreasonable fears concerning addiction, and poor ed-

ucation of health professionals on pain management have also been important factors. Legislative reform, improved pharmaceutical management, and training can reduce these blocks to outpatient palliative care.

Summary

In this chapter we have described options for improving the economic efficiency and accessibility of health services through alternatives to the existing organization of hospitals in many countries. At a time when nearly every country in the world is experiencing severe problems in the provision of an adequate level of services to its population, the implications of the demographic and epidemiological transitions lend added urgency to the need for changes in the manner in which health services are delivered. The rise of the AIDS epidemic has placed a further stress on health systems that may be bankrupted if no changes are made in the technologies and settings used to deliver services. There is much scope for improving the cost-effectiveness of overall service delivery through a reorganization of the functions of specific types of facilities and a reallocation of resources across these facilities. The alternatives seek to match the need for specific services to the least-cost provider of these services. As a result, the case mix of various levels of facilities should become more consistent with their technical capacity, with fewer relatively simple cases being treated in higher-level hospitals. A more precise definition of the roles of each type of facility in the referral system is a necessary first step toward a service delivery network that is more cost-effective.

On the demand side, further steps should be made to encourage users of health services to follow the desired referral pattern. Such steps include improving the reliability and quality of services in lower-level facilities in a way that will be perceived by consumers and structuring prices to provide an incentive to enter the network at the most appropriate facility. Investing in improved quality and reliability at less technologically complex facilities will be costly, and it is an empirical question as to whether the benefits will outweigh the costs. In the absence of empirical data, however, it is presumed that the reductions in overcrowding in hospital outpatient departments and the reduced travel and time costs of consumers who could now rely on the nearest facility for most of their service needs would improve economic efficiency and access.

On the supply side, administrative and financial functions of the management of primary and secondary care facilities that are integrated at the district level (the district primary health care complex) yield better linkages in logistics, training, and patient care than do separate manage-

ment functions at each level of care. Effective integration requires considerable management skills, however, and some separation may be necessary in countries in which lack of a sufficient cadre of trained health system managers might lead to inadequate support for primary health care. Nevertheless, coordination, whether through formal administrative integration or other means, such as the establishment of district health management teams or regional health boards, is necessary to achieve efficient service delivery across levels of care.

In addition to generating a more efficient pattern of use for existing hospitals, alternative treatment strategies and specialized facilities that provide a limited number of specific services can also yield improvements in efficiency and access. Some examples of changes in treatment protocols based on new technological approaches that have been proven to be cost-effective in some developing countries include greater reliance on ambulatory services (for example, ambulatory surgery), new pharmaceuticals for tropical and communicable diseases, and the use of urban maternity centers for low-risk deliveries. Changing the function of underused hospitals to extended care facilities or creating other alternatives for persons who do not require the same level of skilled nursing care that is available in hospitals can also lead to systemwide improvements by freeing hospital beds for those with acute conditions. Finally, the use of alternatives to inpatient care for palliation of chronic diseases such as cancer can be effective, humane, and less expensive than hospitalization.

Appendix 6A. Quality and the Market for Health Services

This appendix outlines the role of quality in the market for health services. The issue of quality enters into both demand- and supply-side decisions. In a competitive market, quality plays an equilibrating role that parallels and interacts with that of price. The important differences are that the signs of the demand and supply responses are reversed compared with those of price and that, unlike price, quality enters directly into the cost function. On the demand side, an increase in quality increases the quantity of services demanded. On the supply side, an increase in quality is accompanied by an increase in cost that reduces the quantity of services supplied at any given price. The market equilibrium will occur at the point at which the price and quality elicit a demand that is equal to supply.

For services produced by public hospitals, the price (fee) and quality are both policy choices made by the government producers. The quantity of services is determined by the reaction of consumers to the price and

quality of services offered if, as is commonly the case in lower-level hospitals, demand is below maximum capacity. A simple analytical framework is derived below to describe the market for services produced by public hospitals. The framework is used to examine the change in quantity demanded when fees are increased and the new revenues are used to improve quality.[2] The quantity of services demanded (Q) is specified as a function (D) of price (P) and quality (Z),

$$Q = D\,(P,Z)\,.$$

The cost (C) of supplying this quantity of services is specified as a function of quantity and quality,

$$C = C\,(Q,Z)$$

or, substituting the demand function for Q,

$$C = C\,[D\,(P,Z), Z]\,.$$

Fiscal solvency requires a revenue balance equation specifying that total revenues, including fee revenue ($P \cdot Q$) and subsidies (S), must equal costs,

$$C\,[D\,(P,Z), Z] = P \cdot D(P,Z) + S\,.$$

Totally differentiating the revenue balance equation allows an examination of the effects of changes in prices and quality,

$$[C_Z + C_Q \cdot D_Z - P \cdot D_Z]\,dZ = [P \cdot D_P + Q - C_Q \cdot D_P]\,dP\,,$$

which is written using the convention that subscripts indicate partial derivatives (for instance, $C_Z = \partial C/\partial Z$) and noting that the subsidy is fixed. This result can be recast in terms of elasticities as:

$$E_{ZP} = \frac{dZ}{dP} \cdot \frac{P}{Z} = \frac{(1 + \eta_{QP} - \eta_{CQ}\,\eta_{QP})}{(\eta_{CZ} + \eta_{CQ}\,\eta_{QZ} - \eta_{QZ})} \tag{6-1}$$

where E_{ZP} is the expression for the total elasticity of quality with respect to a change in prices that retains the equality of total revenue and total cost and where η_{CQ} and η_{CZ} are the partial elasticities of cost with respect to quantity and quality and η_{QP} and η_{QZ} are the partial elasticities of quantity with respect to price and quality.[3]

The total elasticity of quantity with respect to price can be found by totally differentiating the demand function and rewriting the result in terms of elasticities:

(6-2) $$E_{QP} = \eta_{QP} + \eta_{QZ} \cdot E_{ZP}.$$

The total response of quantity to price is thus decomposed into two components, giving the response to the price change itself and the response to the quality change made possible by the increased revenues. The total response will be positive if the quality term is sufficiently great to offset the negative price term.

Some idea of the total response of quantity to price can be obtained by substituting plausible estimates of the separate elasticities comprising E_{ZP} and then E_{QP}. Empirical application of the formulas requires the use of the same definition and scale for quality in both the demand and supply functions. In the available empirical studies of quality and demand carried out thus far, quality has been associated with a productive input such as supplies or pharmaceuticals. Using this definition, we can roughly estimate that the partial elasticity of demand response to a change in drugs, which we will substitute for η_{QZ}, is in the neighborhood (Mwabu and Mwangi 1986) of 0.3. An estimate of the partial elasticity of demand with respect to price is approximately -0.05 (Akin and others 1991). To turn to cost elasticities, if we assume the underlying production function is linear homogenous, the elasticity of cost with respect to output is 1. Finally, if drugs and supplies are 40 percent of cost, an estimate of the elasticity of cost with respect to supplies, which we will substitute for η_{CZ}, is 0.4.

Substituting these rough estimates of the underlying elasticities into equation 6-1 gives an estimate of about 2.5 for E_{ZP}. Making further substitutions in equation 6-2 yields an estimate of about 0.6 for the elasticity, E_{QP}, giving the total response in the quantity demanded to an increase in fees. Thus, if plausible estimates of the elasticities are used, it is concluded that imposing user fees to improve quality is justified on the basis of willingness to pay.

Given the lack of strong empirical estimates of the elasticities it is appropriate to examine the sensitivity of E_{QP} to alternative values. The elasticity of demand and cost with respect to quality are especially roughly estimated. For the specified values of η_{QP} and η_{CZ}, table A6-1 gives the critical value of the elasticity of demand with respect to quality,

Table A6-1. Critical Value[a] *of* η_{QZ}

η_{QP}	$\eta_{CZ} = 0.1$	$\eta_{CZ} = 0.5$
-0.05	0.01	0.03
-0.50	0.05	0.25

a. Below this value the total response of quantity to a price increase is negative.

below which the total response to a price increase is negative. If the price elasticity of demand is high, say, above -0.5, the positive sign for E_{QP} is sensitive to the size of the quality elasticities. However, at the low price elasticities suggested thus far by the econometric research of Akin and others (1991) and Gertler and van der Gaag (1988), the positive value for E_{QP} is robust, and the policy conclusion is that the imposition of fees is warranted by the demand for increased service quality.

Notes

1. This section draws from Barnum and Greenberg (1991).

2. The framework uses, as a point of departure, the supply and demand system specified by Jimenez (1987, appendix C). The authors appreciate a discussion with Jeffrey Hammer of the material in this appendix.

3. For example, $\eta_{CZ} = (\partial C/\partial Z)/(C/Z)$ is the partial elasticity of cost with respect to quality.

7. Conclusions

Our goal in this book has been to contribute to the development of policies that will improve the efficiency of the use by the health sector, and especially by hospitals, of public sector resources. Hospitals have held a pariah status in the health policy dialogue because of the amount of recurrent resources they absorb and the recognition that other types of interventions are much more cost-effective ways to provide the services most needed in developing countries. Yet, because they account for such a large share of public sector health expenditure, improvements in the efficiency of resource use by hospitals can yield tremendous benefits for the entire sector. To assess hospital resource use in the public sector, we have focused on the broad issues of resource allocation, costs, and financing. Improving the economic efficiency of the health sector entails improving resource allocation between hospital and nonhospital activities and between hospitals at different levels of the health system. Knowledge of hospital costs, particularly the average and marginal costs of service provision, is needed to analyze the technical efficiency of hospitals and highlight areas in which waste can be reduced. A keenly debated policy area concerns hospital financing, particularly the issue of recovering costs through user charges and health insurance. Policy in this area is also important to the efficiency and equity of resource use in that financial incentives will affect the behavior of health care providers and consumers. This study of the costs, financing, and use of public sector resources by hospitals will, it is hoped, assist developing countries in their search for solutions to the problems of economic inefficiency and financial crisis in the health sector. The principal conclusions are recapitulated below, followed by suggestions for further research.

Hospital Resource Use

Most governments of developing countries spend more than half of their recurrent health resources on hospitals. This does not necessarily imply a misallocation of resources. Hospitals are complex, high-cost institu-

tions in relation to primary health care facilities, and their very existence means that they will absorb a considerable amount of the total health budget, especially in small countries. Moreover, the public hospital share does not, by itself, demonstrate the level of commitment or the effectiveness of a country's PHC strategy. The health gains achieved by countries such as China and Sri Lanka suggest that the way a country spends its health resources on primary care is more important than simply the percentage of the total budget allocated to nonhospital activities. Nevertheless, it is clear that in some countries, the consumption of recurrent budgets by hospitals, especially tertiary hospitals, is starving other activities. In chapter 2 we provided a considerable array of evidence to support the frequent contention that, especially in poor countries, nonhospital interventions are generally more cost-effective than those performed in hospitals. Thus, it would be wise for policymakers to ensure the recurrent resource requirements of priority (that is, preventive and promotive) health programs. This recommendation holds at all times but is especially relevant during periods of acute fiscal constraints, when maintaining this commitment would likely entail a reallocation of resources away from hospitals.

The need to reallocate resources from more complex to more basic levels of the health care system is a recurring conclusion, discussed in chapters 2, 3, and 6. The necessity derives from the need to improve the quality of services at basic facilities in order to encourage consumers to use local, low-cost providers instead of hospital outpatient departments. If the referral system worked as intended, a greater number of patients could be served by national health systems at less cost to consumers. Instead, the situation often found in developing countries is one of low use of basic facilities because the public perceives the services at these levels to be of poor quality or unavailable, coupled with overcrowded tertiary care facilities, often with patients who do not require complex treatments. A priority for improving the economic efficiency of the health sector, therefore, is to improve quality and the perception of quality at lower-level health facilities. The suggested means for doing so is the reallocation of recurrent resources from tertiary hospitals to primary health care and first-level referral facilities.

Without substantial changes in the way services are currently organized and delivered in most developing countries, the financial crisis in the health sector is likely to worsen. It was demonstrated in chapter 2 that older persons use hospitals disproportionately, and the average cost of treating noncommunicable diseases is greater than for conditions that are of greater consequence for younger age groups. Therefore, as populations age and noncommunicable diseases become relatively more prevalent, the resource requirements of hospitals will grow if there is no change in modes of service delivery. For countries on the brink of this

transition, there is a need to reallocate resources from treatment in hospitals to prevention and control of chronic diseases in order to avert or minimize future high hospital costs. Alternatives for treatment and palliation in lower-cost settings should also be assessed and used in situations in which they are determined to be cost-effective. Improvements in referral and the development of alternative treatment settings should seek to match patient needs with the least-cost provider of services.

Another reason for consideration of nonhospital treatment settings and other forms of technological alternatives to the hospital model adopted from industrialized countries is that the relative prices, or scarcity, of such inputs as staff, equipment, supplies, and maintenance in developing countries differ greatly from those in richer countries. In general, the share of labor in hospital recurrent costs is lower in poorer countries because imported inputs are expensive in relation to labor, and there is apparently only limited substitution of labor for supplies. The higher nonlabor share does not imply that drugs and supplies are available in sufficient quantities. And in many poor countries, the share of personnel is very high (more than 60 percent of hospital recurrent costs), reflecting a low-productivity environment in which shortages of supplies and inadequate maintenance are commonplace. These circumstances demand the assessment of alternative technologies to allow for the provision of services in countries in which skills and imported inputs are in short supply.

The limited available evidence suggests that public expenditure on hospitals is biased in favor of urban populations, which tend to have higher incomes than rural populations. An urban bias in hospital use is expected, given the economic logic behind locating hospitals in urban areas. Excessive concentration of resources in hospitals, however, limits the availability of quality services in rural areas and thus hinders the functioning of the referral system. Reallocation of resources in favor of basic services would not only increase the efficiency of the health sector but would improve equity by providing poorer persons who live in rural areas greater access to health care.

Hospital Costs

Variation in cost estimates results not only from the relative efficiency of input use but also from variation in case mix and quality, because studies of the average or marginal cost of hospital services do not typically adjust for the clinical characteristics of the hospital's patients (that is, the case mix) or the quality of services provided. This limits but does not exclude the use of cost studies in evaluating hospital performance. The growing number of step down accounting studies demonstrates the feasibility of

conducting this type of average cost analysis, and despite the inability to adjust fully for quality and case mix, the studies have yielded conclusions that are relevant to policy, especially when the estimates are interpreted in the context of hospital service statistics. One important finding is that the budgetary information available centrally or even at the hospital level typically understates the actual resources that can be spent by or on behalf of the hospital. This suggests that financial information systems could be made more transparent to give managers a clear idea of the resources with which they are working and, furthermore, how these resources are being used in specific departments of the hospital.

The magnitude of the difference in average costs found between tertiary and lower-level hospitals suggests that there is considerable scope for savings (or increased service provision) by shifting patients who do not require sophisticated care from facilities that are more complex to those that are less so. Moreover, basic hospitals often exhibit lower bed turnover and occupancy rates than upper-level hospitals, which reflects a lack of demand for these less expensive facilities. This lends further weight to the policy conclusion that efforts should be made to encourage patients to use their local facilities for basic care. Although the suggested means to accomplish this—reallocating resources from upper- to lower-level facilities to improve their quality—will increase their total recurrent cost, demand response to the improvement in quality may be sufficient to lower average cost per patient. In any event, the expected change in demand patterns should yield improved economic efficiency throughout the network of providers with the result that a greater number of patients can be served without an increase in the level of resources.

Statistical studies of hospital costs in developing countries are in a nascent stage, and thus generalization of the results can only be made cautiously. The identification and analysis of hospital cost functions represent an important development because of the answers these studies could provide to some important policy questions, such as the relation between marginal and average costs and the optimal size and scope of hospitals. For example, one finding from the survey of cost functions in developing countries that is consistent with studies from industrial countries is the absence of long-term economies of scale. The discussion of statistical cost studies in chapter 3 focused on the importance of understanding the nature of the underlying cost function. Although accounting studies probably provide a good approximation of costs in the short run, they may be a misleading foundation on which to base longer-term investment decisions because the actual cost function is likely to differ markedly from the "passive" accounting estimate if substantial changes occur in current levels of output.

When used together, inpatient performance indicators—bed occupancy rate, bed turnover rate, and average length of stay—provide a useful tool for comparing the relative performance of similar types (that is, similar function and case mix) of hospitals within a country. They enable rapid identification (followed, it is hoped, by investigation) of hospitals that exhibit signs of technical inefficiency, such as a basic hospital whose patients have long average stays. Similar indicators are needed to evaluate outpatient performance. Although rectifying technical inefficiency typically requires the skills of managers rather than economists, widespread problems in the use of personnel and technical inputs are often related to the presence of inappropriate incentives generated by the health financing system. Such deficiencies in incentives should be addressed by economic analysis.

Hospital Financing

Government intervention in the market for hospital (and other health) services is necessary because substantial market failures are associated with this sector. In other words, some of the assumptions underlying the optimality of free market activities, such as the availability of sufficient information about the nature of the services for providers and consumers to make informed decisions, are violated. As a result, the activities of unregulated private buyers and sellers do not yield an economically efficient allocation of resources.

The desire to explore alternatives to central treasury financing of the health services has typically been motivated by a desire to expand the availability of resources, but any program of health financing has implications for the efficiency and equity of the sector in addition to its effects on revenues. The review of the experience of various countries and analysis of market failures indicate that there is no single, optimal solution to the question of how to finance hospitals, or more generally, the health services. All the options—ranging from public financing and provision, to modest user charges, to explicit forms of health insurance—have potentially undesirable effects on resource allocation. Nevertheless, in all countries there is scope for making practical improvements to existing financing systems by improving incentives to support allocational efficiency.

In chapter 4 principles were elaborated for setting fees in public sector facilities that attempt to balance the objectives of efficiency, equity, and revenue collection. Hospital prices should (a) not limit access of poor persons to essential services, (b) provide proper signals for the use of health resources across services and facilities, (c) be linked to the quality of services, and (d) be subsidized for services that would otherwise not be adequately consumed as a result of market failures. A guide for setting

prices that is consistent with these principles is called optimal pricing. Optimal prices reflect a combination of the demand characteristics (that is, income and price elasticities) associated with a specific service, distributional goals, and the cost of providing the service. Prices are set lower for services that are consumed disproportionately by the poor yet contribute to efficiency goals by interfering minimally with private preferences.

The review of country experience with user charges presented in chapter 5 found that fees usually contribute only a small proportion of total expenditures on public hospitals. Yet the countries that are exceptions to this finding demonstrate the feasibility of cost recovery to provide substantial revenues in both low- and middle-income economies. Experience suggests some complementary recommendations to the pricing principles summarized above. First and foremost, regular adjustment of fee levels to keep pace with inflation should be institutionalized as part of any fee system, perhaps by tying charges to price or cost indexes. To the extent possible, price changes should be an administrative or technical decision rather than a political one. Second, the setting of prices should be decentralized in countries that are geographically large or in which there is a considerable degree of regional variation in income levels. Third, price schedules should be kept simple (for example, a single charge for each broad category of services rather than multiple prices for fine gradations of services within a category) if cost-recovery goals are modest and institutional development is weak. A fourth recommendation that derives from analysis of the incentives that should be in place but for which more empirical support is needed is that hospitals should be allowed some degree of fee retention and that they should use the revenues to purchase recurrent inputs needed to enhance the quality of services. A fifth recommendation that is related to equity rather than revenue is that no personal criteria for exemptions from payment other than patient income should be used. However, patients who use services with important externalities, such as treatment of certain communicable diseases, should be exempted from payment so that the use of such services will be encouraged.

In most countries, fees are set at minimal levels and thus do not alter financial incentives very much from those prevailing under completely free systems of service provision. Where fee levels are low, governments are implicitly providing financial protection against the risk of a high-cost hospitalization. Given the expensive nature of inpatient care in relation to the incomes of both poor and many nonpoor persons, substantial cost recovery for inpatient services requires some type of explicit risk-sharing mechanism.

Health insurance programs vary along a number of dimensions, including the hospital payment mechanism, the services covered and

excluded, the role of the insuring institution, the extent of cost sharing required of beneficiaries, and the proportion of the population covered by the program. The specific institutional arrangements have different implications for client and provider behavior. As with the principles for user charges (but more important, given the stronger effects of the price signals for substantial reimbursement of inpatient costs), health insurance programs should be designed with an eye toward providing the proper incentives to providers and their clients for efficient and equitable behavior.

Experience suggests that institutions that administer insurance programs should not serve merely as financial conduits between beneficiaries and providers. They should actively encourage (via incentives or some type of contractual relationship) providers and consumers to behave in a manner that limits cost escalation, excessive allocation of resources to curative care, use of inappropriate medical technology, and inadequate access for financially disadvantaged persons. Incentives to providers, such as those inherent in prepaid capitation programs, seem to be more effective in limiting cost escalation than do cost-sharing obligations directed at consumers. The underlying reason for this is a critical market failure: asymmetry in the availability of information on the nature of health services between doctors and their patients. This asymmetry typically results in patients' relying on their provider for information on the services they need. One way to address this conflict of interest for providers is to make the insuring institution the "agent" for the patient, thus removing the decision on the course of treatment from the provider to be reimbursed.

Health financing policy should be coordinated across levels of care on a regional or national basis. Health facility user charges should reflect incentives to use the least costly provider and should reward (penalize) appropriate (inappropriate) use of the referral network. Likewise, denying reimbursement for patients who self-refer to hospitals is an important institutional measure that insurance programs should use to support economically efficient use of a country's health facilities. In other words, health financing policy should complement measures to reallocate resources from tertiary care and improve the quality of basic health services. Together, these policies can lead to improved efficiency and availability of services throughout the health system.

Suggestions for Future Research

Because of data limitations and the need to restrict the scope of the inquiry, the answers to many questions raised in this book remain tentative. We have attempted to extract the salient implications of observations and experiences from a wide cross-section of countries, using

economic analysis to sharpen the focus on specific issues. Such a general survey can flag broad issues and outline potential solutions for consideration by those whose task it is to develop a strategy for hospitals and the health sector in a specific country. Defining specific policies, however, requires further research on many detailed questions of resource allocation, costs, and finance in the countries concerned. Some of the most important questions and possible directions for refinement are noted below.

Resource Allocation

Identification of the optimal amount of resources to allocate to hospitals under alternative resource constraints and at different levels of development of the health sector remains a significant problem for health system planning. This identification requires clear definitions of hospitals (at all levels) and their role in a country's health system. Given these definitions, an understanding of both the efficiency and equity implications of hospital resource use can enable policymakers to determine the appropriate share of public sector resources to devote to hospitals.

With regard to economic efficiency of resource allocation in the health sector, a fundamental need is to identify the health outcomes of hospital services. To do this, better information on inpatient and outpatient case mix is necessary, especially by level and ownership of hospital. Given the case mix, information is needed on the effect of hospital services on health status for individual case types and on the cost and effect of alternative nonhospital interventions. Measurement of health status on admission and discharge and attribution of the change in health status to hospital intervention would be an ambitious research effort. Construction of scales from physiological measures and the application of these scales to estimate the effect of hospital services would require careful statistical design and training of evaluators. One possible direction for such research is that taken by Watters and others (1989), who adapted the strategy used by Knaus and his colleagues (1985; later updated; see Zimmerman 1989) to classify critically ill patients at Zambia's University Teaching Hospital into severity categories, allowing for a case mix–adjusted assessment of the effects of alternative types of treatments. Although such studies would be expensive and not routinely practical, the information gained would provide insights that not only could answer economic questions related to the efficiency of resource allocation within and across hospitals but would aid management and planning of the health sector more generally. Better understanding of the outcomes associated with health services would improve the interpretation of process measures (for example, by enabling the comparison of observed input use with normative models associated with desired outcomes),

which are more readily available to management than outcome measures. It would also help clarify the role of hospitals, especially hospitals of different levels, in planning the health system.

The reallocation of resources from more complex to less complex facilities as a way of expanding the availability of services (given a fixed budget) at the national or regional level requires an assessment of the strengths and weaknesses of the existing referral system. For this purpose analytical tools need to be developed. An important focus of applied research in this area would be analysis of the costs, case mixes, and outcomes of services provided in different settings that could be substituted for each other. A likely example would be a study of hospital outpatient and health center services in the same district. The analytical method used by Shepard and others (1991) to compare the costs and outcomes of hernia surgery in the University Hospital and "outpatient hospitals" in Cali, Colombia, provides a good model for such research.

With regard to equity, more research is required on the distribution of benefits from hospital services by income, region, and age. Also, equity of hospital use by sex should be studied. More detailed information would be useful on distributional equity with regard to access to, use of, and benefit from hospital services in relation to the need for services. Studies should also distinguish between use of public and private services (for specific types of services used) by different socioeconomic groups, and the equity goal should be clearly defined. For example, the goal could be defined as equal per capita use of hospitals for urban and rural populations or as equal access to necessary care. The latter is a broader goal that acknowledges greater use of hospital (especially outpatient) services by the urban population. In the pursuit of such an objective the focus would be on improving the referral system to increase the likelihood that rural persons would gain access to hospital care when needed. Assessing use in relation to "need" would be much more difficult, however, than simply measuring differences in per capita use across income and geographic groups. Yet the available information even on this latter issue is scanty, although hospital resources appear to be used disproportionately by middle- and upper-income groups because of their largely urban location. More likely is that there is considerable variation across countries and regions in the equity of hospital resource use depending on prevailing economic structure, population distribution, and social norms.

Hospital Cost and Production

We need a better understanding of the institutional structure of hospitals and the maximizing behavior of hospital managers in order to predict their response (in terms of hospital service mix and technology choice)

to alternative policies. What are the objectives of hospital managers? How does the behavior of physicians serving both public and private patients affect service delivery? How does physician education affect the choice of technology (including the use of hospital as opposed to non-hospital services)?

Improved understanding of the production process and cost functions facing hospitals could result from a clearer knowledge of the optimizing behavior (or lack thereof) of hospital managers. We are not aware of any published studies of statistical production functions for developing countries, and, as reviewed in chapter 3, only a handful of statistical cost functions have been estimated. Better understanding is needed of the hospital production process. Statistical production functions could add insight into the degree of substitutability among hospital inputs and the role of technological change. Knowledge of cost functions also helps in assessments of marginal costs and the importance of economies of scale. Hospital service statistics should be more fully integrated into this analysis, and to this end there is a need to develop useful indicators for outpatient performance similar to those available for inpatient services.

The problem of defining and identifying quality of hospital services pervades all the research needs discussed in this section, whether concerning resource allocation, cost and efficiency, or financing. Identifying the dimensions of quality enhancement and then measuring the recurrent costs of improving quality need further study. Ideally, quality should be defined in terms of health outcomes, but this is probably unrealistic except for special studies that specifically measure health status. Alternative indicators of quality can be developed, such as the use of benchmark protocols for inputs (for example, the level and type of training received by staff, the quantity and selection of drugs, and the kind of diagnostic equipment available and its state of repair), and case mix–adjusted rates of infection, complication, and survival. Calibration of these alternative normative indicators of quality can be based on research that uses special surveys to link process and outcome measures. Such calibration would make it possible to use reasonable inferences to monitor quality on an ongoing basis.

Very little is known about the crucial behavioral parameters needed to understand the role of quality as an equilibrating force in health markets. The relation of quality as perceived from the point of view of hospital users to quality as actually measured by health outcome has not been adequately examined. Research is needed into how consumers perceive quality, the response of demand to changes in quality, the effect of quality changes on the cost of services and, finally, the relation between prices and quality. A potential framework for research on these questions is provided in the appendix to chapter 6. Data to examine these questions can be accumulated by linking specific information about

health service providers (including hospitals) to data from household surveys.

Financing

In order to achieve equity, efficiency, and revenue goals, more needs to be known about the effect of prices on household consumption choices. The econometric studies cited in chapter 4 provide evidence on the price elasticity of hospital services but do not identify the effect of price changes on the pattern of consumer expenditures for other goods. Although the price elasticity of hospital services is low, increased prices for hospital services may bring about reductions in the consumption of other basic items, such as education or food, and thus have a substantial effect on welfare. Estimation of demand and income elasticities within the framework of a full household model is needed to understand the cross-elasticities associated with an appraisal of the effect of hospital fees on welfare. In addition, the household model must be extended to deal with changes in asset holdings if the mechanisms that households use to cope financially with the sporadic high cost of hospital care are to be understood. This applies to costs actually incurred by households and to estimates of the costs that would be incurred if the subsidy to public hospital care was reduced or eliminated.

We single out hospital services for special treatment, rather than merely using an aggregate category of "health care" expenditures because hospital care, when it is required, is by far the largest health expenditure item for households or the health system. Careful design of future surveys to address the question of hospital pricing may provide information on the best ways to bundle hospital services. Extension of models of household utility maximization to encompass separate consumption categories for hospital services and to include changes in asset holding provides a substantial research challenge both in obtaining health-specific household consumption data and in the design of appropriate econometric models. However, work by Gertler and van der Gaag (1990), Pitt and Rosenzweig (1990), Rosenzweig and Schultz (1983), and Behrman and Deolalikar (1990), which adds nutrition and health to the household utility framework, is moving in the appropriate direction and indicates that the challenge can potentially be met.

The equity implications of alternative insurance schemes could be illuminated by research on the kind and level of services demanded by different income groups according to insurance status. By linking information from client surveys with studies of the resource content of hospital services it would be possible to reconstruct the actual distribution of the benefits from hospital resources used. Critical questions concern the resource benefits flowing to insured population subgroups,

such as civil servants and high-income urban workers, compared with those received by rural or low-income urban populations. This question can be extended, by using information on tax incidence and resources received from cost recovery, to compare the incidence of benefits with the incidence of payments.

A better understanding of the incentive effects of alternative plans on producers and clients is needed. There is an expectation that, given similar case mix, there will be lower admission rates under prepaid plans than under retrospective reimbursement, and producers will use fewer resources per admission under case-based payment than they would under fee-for-service payment. There is also an expectation that the insured will use more services than the uninsured. Difficult but critical research problems are posed by the need to test whether these expectations are valid, and if valid, to estimate the extent to which the changed use of services conflicts with health needs. More research is also needed on the institutional features complementary to reimbursement mechanisms that can help contain costs and mitigate the potentially deleterious effects of insurance programs on equity, resource allocation, and technology choice.

Finally, this book has been restricted primarily to a study of public hospitals. One important financing alternative that was not examined was the possibility of explicitly shifting part of the burden for delivering hospital services from the public to the private (voluntary and for-profit) sector. Consideration of this possibility requires answers to questions about the relative efficiency of public and private hospitals as determined by comparative cost functions or accounting cost studies. Setting up appropriate efficiency studies is difficult because differences in quality and case mix, which may be substantial, must be measured. The incorporation of private hospitals in health care planning also raises important questions related to the setting of fees in the public sector. How should access for the poor be guaranteed? Should private physicians be charged for the use of public facilities? Competition between public and private hospitals might affect the availability of skilled personnel in public facilities. To what extent would this affect the quality of services available to poor people? What are the cross-elasticities between prices for public and private services? How is the use of public and private services affected by income, distance between hospitals, and differences in quality? These important questions are beyond the scope of the present study of resource use, cost, and financing of hospitals in the public sector. Nevertheless, they remain important policy issues that governments and researchers must address.

References

ADB (Asian Development Bank, World Bank Economic Development Institute and East West Center). 1987. *Health Care Financing*. Washington: EDI/World Bank.

AHA (American Hospital Association). 1989. *Hospital Statistics*. Annual. Chicago.

Aigner, Dennis J., C. A. K. Lovell, and Peter J. Schmidt. 1977. "Formulation and Estimation of Stochastic Frontier Production Models." *Journal of Econometrics* 6(1):21-38.

Akin, John. 1987. "Health Insurance in the Developing Countries: Prospects for Risk Sharing." In Asian Development Bank, World Bank Economic Development Institute and East West Center, *Health Care Financing*. Washington, D.C.: EDI/World Bank.

Akin, John, Nancy Birdsall, and David de Ferranti. 1987. *Financing Health Services in Developing Countries: An Agenda for Reform*. A World Bank Policy Study. Washington, D.C.

Akin, John, Charles C. Griffin, David Guilkey, and Barry M. Popkin. 1986. "The Demand for Adult Outpatient Services in the Bicol Region of the Philippines." *Social Science and Medicine* 22(3):321–28

Akin, John S., Hazel Denton, David K. Guilkey, Ronald J. Vogel, and Annemarie Wouters. 1991. "Health Care Costs, Demand, and Cost Recovery in Ogun State, Nigeria." World Bank, Western Africa Department, Population and Human Resources Division, Washington, D.C.

Alailima, P., and F. Mohideen. 1984. "Health Sector Expenditure Flows in Sri Lanka." *World Health Statistics Quarterly* 37(4):403–20.

Anderson, David L. 1980. "A Statistical Cost Function Study of Public General Hospitals in Kenya." *Journal of Developing Areas* 14(2):223–35.

APH (Association Pour la Promotion de l'Hôpital—Saint-Louis). 1987. *Règlement Intérieur*. Saint-Louis, Senegal.

Arrow, Kenneth. 1976. "Welfare Analysis of Changes in Health Coinsurance Rates." In Richard N. Rosett, ed., *The Role of Health Insurance in the Health Services Sector*. New York: National Bureau of Economic Research.

Atkinson, Anthony B., and Joseph E. Stiglitz. 1980. *Lectures on Public Economics*. New York: McGraw-Hill.

Baker, W. G. 1977. "In-depth Study on Hospital Expenditure." Assignment Report. WHO Regional Office for the Western Pacific, Manila.

Banta, David. 1987. Background Report for World Bank health sector mission to China. World Bank, East Asia and Pacific, Country Department II, Environment, Human Resources, and Urban Development Operations Division, Washington, D.C.

Barlow, Robin. 1976. "Applications of a Health Planning Model in Morocco." *International Journal of Health Services* 6(1):103–22.

Barlow, Robin, and Lisa M. Grobar. 1986. "Cost and Benefits of Controlling Parasitic Diseases." PHN Technical Note 85-17, Population, Health, and Nutrition Department. World Bank, Washington, D.C.

Barnum, Howard. 1986. "Cost Savings from Alternative Treatments for Tuberculosis." *Social Science and Medicine* 23(9):847–50.

——. 1987. "Hospital Expenditure in Indonesia." PHN Technical Note 87-17, Population, Health, and Nutrition Department. World Bank, Washington, D.C.

——. 1989. "Economic Issues in Planning the Use of Health Resources for Chronic Diseases in China." World Bank, East Asia and Pacific Country Department II, Environment, Human Resources, and Urban Development Operations Division, Washington, D.C.

Barnum, Howard, Robin Barlow, Luis Fajardo, and Alberto Pradilla. 1980. *A Resource Allocation Model for Child Survival.* Cambridge, Mass.: Oelgeschlager, Gunn and Hain.

Barnum, Howard, and Robert Greenberg. 1991. "Cancers." In *The World Bank Health Sector Priorities Review.* Paper HSPR-01. World Bank, Population and Human Resources Department, Population, Health and Nutrition Division, Washington, D.C.

Barnum, H. N., D. Tarantola, and I. F. Setiady. 1980. "Cost-Effectiveness of an Immunization Programme in Indonesia." *Bulletin of the World Health Organization* 58(3):499–503.

Baumol, William J., and David F. Bradford. 1970. "Optimal Departures from Marginal Cost Pricing." *American Economic Review* 60(3):265–83.

Behrman, Jere R., and Anil B. Deolalikar. 1990. "The Intrahousehold Demand for Nutrients in Rural South India: Individual Estimates, Fixed Effects, and Permanent Income." *Journal of Human Resources* 25(4):665–96.

Bennis, J. P. Dumont, N. Guedira, L. Jaidi, M. Lahlou, N. E. Ottmani, J. R. Siegfied, and A. Tibouti. 1990. "Etude du Système de Financement des Secteurs de Santé." Report prepared by Groupement ICONE-SEDES (Ingénierie et Conseil en Economie-Société d'Etudes pour le Développement Economique et Social). Report prepared for the Ministry of Public Health of the Kingdom of Morocco, Rabat.

Bitran, Ricardo, Munkatu Mpese, Sebinwa Bavugabgose, Muanza Kasonga, Nsembani Nsuka, Taryn Vian, and Kongoba Wambenge. 1987. "Zaire Health Zones Financing Study. June–October 1986." Resources for Child Health (REACH) Project, USAID. Arlington, Va.

Bitran-Dicowsky, Ricardo, and David W. Dunlop. 1989. "The Determinants of Hospital Costs: An Analysis of Ethiopia." PPR Working Paper 249. World Bank, Population and Human Resources Department, Population, Health, and Nutrition Division, Washington, D.C.

Blanpain, Jan E. 1987. "Hospital Facilities Issues in Chronic Disease Management." World Bank, East Asia and Pacific Country Department II, Environment, Human Resources, and Urban Development Operations Division, Washington D.C.

Bloom, Gerald, and Caroline Temple-Bird. 1988. "Medical Equipment in Sub-Saharan Africa: A Framework for Policy Formulation." Institute for Development Studies (IDS) Research Report Rr19. Sussex, England.

BMOH (Botswana, Ministry of Health). 1979. *The Financing of Health Services and Activities in Botswana.* Second Study. Gaborone.

Bobadilla, José-Luis, and Godfrey J. A. Walker. 1991. "Early Neonatal Mortality and Cesarian Delivery in Mexico City." *American Journal of Obstetrics and Gynecology* 164(1):22–28.

Boiteux, M. 1971. "On the Management of Public Monopolies Subject to Budgetary Constraints." *Journal of Economic Theory* 3(3):219–40.

Bos, Dieter. 1985. "Public Sector Pricing." In Alan J. Auerbach and Martin Feldstein, eds., *Handbook of Public Economics.* Vol. 1. Amsterdam: North-Holland.

——. 1986. *Public Enterprise Economics: Theory and Application.* Amsterdam: North-Holland.

Breyer, Friedrich. 1987. "The Specification of a Hospital Cost Function: A Comment on the Recent Literature." *Journal of Health Economics* 6(2):89–163.

Briscoe, John. 1990. *Brazil: The New Challenge of Adult Health.* A World Bank Country Study. Washington, D.C.

Budd, Karen. 1989. "Study of Budget Allocation Policy and Programming Issues in Niger, Including an Analysis of Budget Allocations in 1988 (Part 1)." Prepared under USAID Project 683-0254. Tulane/Abt Associates. July. Cambridge, Mass.

Bumgarner, J. Richard, ed. 1992. *China: Long-Term Issues and Options in the Health Transition.* A World Bank Country Study. Washington, D.C.

Carreon, Gabriel G. 1982. "The Role of the Hospital in Primary Health Care." *World Health Forum* 3(2):152–55.

Cavazos-Ortega, Nora, Aurora del Rio-Zolezzi, José Antonio Izazola-Licea, Miguel Angel Lezana-Fernández, and José Luis Valdespino-Gómez. 1989. "Años de Vida Potential Perdidos: Su Utilidad en el Análisis de la Mortalidad en México." *Salud Pública de México* 31(5):610–24.

Centers for Disease Control. 1989. *Morbidity and Mortality Weekly Report (MMWR)* 37(53).

Chen, Jie. 1987. "The Cost of Hospitalization in Shanghai Eighth People's Hospital of China." Shanghai Medical University, Training Center for Health Management.

——. 1988. "The Cost of Hospitalization in Shanghai First People's Hospital, China." Shanghai Medical University, Training Center for Health Management.

Ching, Panfila. 1986. "Public Provision and Demand for Health Services: A Case Study of Bicol." *Philippine Review of Economics and Business* 23(3&4):293–321.

Chomitz, Kenneth. 1986. "Hospitals in Developing Countries: Supply-Side Issues." World Bank, Population and Human Resources Department, Washington, D.C.

CMBT (Christian Medical Board of Tanzania). 1991. Medical Statistics, 1989. Dar es Salaam.

Cohen, Harold A. 1967. "Variations in Costs among Hospitals of Different Sizes." *Southern Economic Journal* 33 (January):355–66.

Collins, David H. 1990. "An Assessment of User Fees in Swaziland's Health Services." Paper prepared for USAID/Swaziland. Management Sciences for Health, Boston.

Cowing, Thomas G., and Alphonse G. Holtmann. 1983. "Multiproduct Short-Run Hospital Cost Functions: Empirical Evidence and Policy Implication from Cross-Section Data." *Southern Economic Journal* 49 (January):637–53.

Cowing, T. G., A. G. Holtmann, and S. Powers. 1983. "Hospital Cost Analysis: A Survey and Evaluation of Recent Studies." *Advances in Health Economics and Health Services Research* 4:257–303.

Cumper, G. C., G. J. A. Walker, and C. MacCormack. 1985. "Evaluation of Health Care in Jamaica." London School of Hygiene and Tropical Medicine, Evaluation and Planning Centre.

Dean, Malcolm, ed. 1981. *The Role of Hospitals in Primary Health Care*. Report of a conference sponsored by the Aga Khan Foundation and the World Health Organization, November 22–26, Karachi, Pakistan.

de Béthune, X., P. Mercenier, and H. Van Balen. 1985. "The Role of the Hospital in Strengthening Primary Health Care: The Experience of Kasongo." Working Paper SHS/EC/85/WPIII.C.5 from the WHO Expert Committee on the Role of Hospitals at the First Referral Level, Geneva, Switzerland.

de Ferranti, David. 1985. *Paying for Health Services in Developing Countries: An Overview*. World Bank Staff Working Paper 721. Washington, D.C.

Djuhari Wirakartakusumah, M., Prijono Tjiptoherijanto, Azwini Kartoyo, Ricardi W. Alibasjah, Darwis Hartono, Eko Ganiarto, and Erinos Muslim Tanjung. 1988. *Evaluasi dan Analisis Biaya Rumah Sakit Fase II* (Phase II Evaluation and Analysis of Hospital Costs). University of Indonesia, School of Economics, Bureau of Planning, Department of Health and the Demographic Institute. Jakarta.

Donaldson, Dayl S., and David W. Dunlop. 1987. "Analysis of Utilization and Financial Information of Ethiopian Government Health Facilities." World Bank, Eastern Africa Department, Population and Human Resources Operations Division, Washington, D.C.

Drummond, Michael F. 1980. *Principles of Economic Appraisal in Health Care*. Oxford: Oxford University Press.

——. 1985. "Survey of Cost Effectiveness and Cost-Benefit Analyses in Industrialized Countries." *World Health Quarterly* 28(4):383–401.

Dunlop, David W. 1973. "The Economics of Uganda's Health Service System: Implications for Health and Economic Planning." Ph.D. diss., Michigan State University.

——. 1984. "Underfinancing of Social Services in Tanzania: The Case of Primary Health Care." Prepared for USAID under contract ORT-0085-0-00-3167-00.

——. 1986. "Issues in Hospital Costs in Poor Countries: Thoughts on a Research Agenda." World Bank, Population and Human Resources Department, Population, Health, and Nutrition Division, Washington, D.C.

Eddy, David M. 1986. "Secondary Prevention of Cancer: An Overview." *Bulletin of the World Health Organization* 64(3):421–29.

Ellis, Randall P. 1987. "The Revenue Generating Potential of User Fees in Kenyan Government Health Facilities." *Social Science and Medicine* 25(9):995–1002.

Epstein, Paul R. and Linda Coultas. 1991. "Kenya: Cost-Recovery Programme." Lancet 337 (June 15):1467–68.

Feachem, Richard G. 1986. "Preventing Diarrhoea: What Are the Policy Options?" *Health Policy and Planning* 1(2):109–17.

Feldstein, Martin S. 1968. *Economic Analysis for Health Service Efficiency*. Chicago: Markham Publishing Company.

——. 1972a. "Distributional Equity and the Optimal Structure of Public Prices." *American Economic Review* 62(1):32-36 and 763 (errata).

——. 1972b. "Equity and Efficiency in Public Sector Pricing: The Optimal Two-Part Tariff." *Quarterly Journal of Economics* 86(2):175–87.

——. 1972c. "The Pricing of Public Intermediate Goods." *Journal of Public Economics* 1(1):45–72.

Fetter, R. B., Y. Shin, J. L. Freeman, R. F. Averill, and J. D. Thompson. 1980. "Case Mix Definition by Diagnosis-Related Groups." *Medical Care* 18 (Supp. 2):1–39.

Fiedler, John L. 1987. "Recurrent Cost and Public Health Care Delivery: The Other War in El Salvador." *Social Science and Medicine* 25(8):867—74.

FKMIS (Federation of Korean Medical Insurance Societies). 1987. *'86 Medical Insurance Statistical Yearbook*. Ninth edition, Seoul.

Forte, P. G. L. 1985. "Costing Inpatient Health Care in Britain: A Review of Methods." *Environment and Planning C: Government and Policy* 3(2):187–98.

Frenk, Julio, José-Luis Bobadilla, Jaime Sepúlveda, and Malaquias López Cervantes. 1989. "Health Transition in Middle-Income Countries: New Challenges for Health Care." *Health Policy and Planning* 4(1):29–39.

Frenk, Julio, Enrique Ruelas, and Avedis Donabedian. 1989. "Hospital Management Staffing and Training Issues." PPR Working Paper 173. World Bank, Population and Human Resources Department, Population, Health, and Nutrition Division, Washington, D.C.

Gertler, Paul, Luis Locay, and Warren Sanderson. 1987. "Are User Fees Regressive?: The Welfare Implications of Health Care Financing Proposals in Peru." *Journal of Econometrics* 36(1/2):67–88.

Gertler, Paul, and Jacques van der Gaag. 1988. "Measuring the Willingness to Pay for Social Services in Developing Countries." Living Standards Measure-

ment Study, Working Paper 45. World Bank, Population and Human Resources Department, Poverty Analysis and Policy Division, Washington, D.C.

——. 1990. *The Willingness to Pay for Medical Care.* Baltimore: Johns Hopkins University Press.

GHAA (Group Health Association of America, Inc.). 1985. *Managed Prepaid Health Care in Latin America and the Caribbean: A Critical Assessment.* Vol. 1. Prepared under Cooperative Agreement LAC 000-A-00-4049-00 for USAID Bureau for Latin America and the Caribbean. Washington, D.C.

GHAT (Ghana Health Assessment Team). 1981. "A Quantitative Method of Assessing the Health Impact of Different Diseases in Less Developed Countries." *International Journal of Epidemiology* 10(1):73–80.

Ghazali, Azizan Aiyub, Haji Abdul Halim Bin Abdul Salam, and Akbar Bin Baba. 1987. "Health Care Financing: Malaysia." Country Paper prepared for the Regional Seminar on Health Care Financing, Asian Development Bank, July, Manila, Philippines.

Ghei, P. N. 1985. "The Role of Hospitals in Primary Health Care (PHC): Report on a Survey of 65 Hospitals in Delhi, 1983." In *Proceedings of International Conference on Hospitals and Primary Health Care.* New Delhi, India.

Gilson, Lucy. 1988. "Government Health Care Charges: Is Equity Being Abandoned?" EPC Publication 15. London School of Hygiene and Tropical Medicine.

Gish, Oscar, Ridwan Malik, and Paramita Sudharto. 1988. "Who Gets What? Utilization of Health Services in Indonesia." *International Journal of Health Planning and Management* 3(3):185–96.

Glewwe, Paul. 1988. "The Distribution of Welfare in Côte d'Ivoire in 1985." Living Standards Measurement Study, Working Paper 29. World Bank, Population and Human Resources Department, Poverty Analysis and Policy Division, Washington, D.C.

Glewwe, Paul, and Kwaku A. Twum-Baah. 1991. "The Distribution of Welfare in Ghana in 1987–88." Living Standards Measurement Study, Working Paper 75. World Bank, Population and Human Resources Department, Poverty Analysis and Policy Division, Washington, D.C.

GOJMOH (Government of Jamaica, Ministry of Health). 1986. Hospital Statistics Report 1985. Health Information Unit. Kingston, Jamaica.

——. 1987. "Hospital Restoration Project." Loan Application Document prepared by the Pan American Health Organization (PAHO), submitted to the Inter-American Development Bank. Kingston, Jamaica.

Government of Ghana. 1985. "Hospital Fees Regulations, 1985." Accra, Ghana.

GOZMOH (Government of Zimbabwe, Ministry of Health). 1991. "Manual for Charges at Government Supported Hospitals." Harare, Zimbabwe.

Grannemann, Thomas W., Randall S. Brown, and Mark V. Pauly. 1986. "Estimating Hospital Costs: A Multiple-Output Analysis." *Journal of Health Economics* 5(2):107–27.

Green, Andrew. 1987. "The Role of Non-Governmental Organizations and the Private Sector in the Provision of Health Care in Developing Countries." *International Journal of Health Planning and Management* 2(1):37–58.

Griffin, Charles C. 1987. "User Charges for Health Care in Principle and Practice." In Asian Development Bank, World Bank Economic Development Institute and East West Center, *Health Care Financing*. Washington, D.C.: EDI/World Bank. Also available as World Bank Economic Development Institute Seminar Paper 37, Washington, D.C.

——. 1989. *Strengthening Health Services in Developing Countries through the Private Sector*. International Finance Corporation Discussion Paper 4. Washington, D.C.

——. 1992. *Health Care in Asia: A Comparative Study of Cost and Financing*. The World Bank: Washington D.C.

Grosh, Margaret E. 1990. *Social Spending in Latin America: The Story of the 1980s*. World Bank Discussion Paper 106. Washington, D.C.

Gsellman, Robert. 1985. "One Hospital's Relationship to a Rural Outreach Program." In *Proceedings of International Conference on Hospitals and Primary Health Care*. New Delhi, India.

Halstead, Scott B., Julia A. Walsh, and Kenneth S. Warren, eds. 1985. *Good Health at Low Cost*. New York: Rockefeller Foundation.

Harris, Jeffrey E. 1977. "The Internal Organization of Hospitals: Some Economic Implications." *Bell Journal of Economics* 8 (Autumn):467–82.

——. 1979. "Pricing Rules for Hospitals." *Bell Journal of Economics* 10 (Spring):224–43.

Hecht, Robert M., ed. 1992. *Zimbabwe: Financing Health Services*. A World Bank Country Study. Washington, D.C.

Hecht, Robert, Catherine Overholt, and Hopkins Holmberg. 1992. "Improving the Implementation of Cost Recovery for Health: Lessons from Zimbabwe." Technical Working Paper No. 2, World Bank, Africa Technical Department, Population, Health, and Nutrition Division, Washington, D.C.

Heller, Peter S. 1975. *Issues in the Costing of Public Sector Outputs: The Public Medical Services of Malaysia*. World Bank Staff Working Paper 207. Washington, D.C.

——. 1979. "Underfinancing of Recurrent Development Costs." *Finance and Development* 16(1):38–41.

——. 1982. "A Model of the Demand for Medical and Health Services in Peninsular Malaysia." *Social Science and Medicine* 16(3):267–84.

Herz, Barbara, and Anthony R. Measham. 1987. *The Safe Motherhood Initiative: Proposals for Action*. World Bank Discussion Paper 9. Washington D.C.

Hijar-Medina, Martha. 1990. "Mortalidad por Lesiones Accidentales E Intencionales en el Distrito Federal de 1970 a 1986." *Salud Pública de México* 32(4):395–404.

Hoare, Geoff, and Anne Mills. 1986. "Paying for the Health Sector." EPC Publication 12. London School of Hygiene and Tropical Medicine.

Hôpital de Saint-Louis. 1986. Financial Report. Ministry of Public Health, Saint-Louis, Senegal.

Horton, Susan, and Pierre Claquin. 1983. "Cost Effectiveness and User Characteristics of Clinic-Based Services for the Treatment of Diarrhoea: A Case Study in Bangladesh." *Social Science and Medicine* 17(11):721–29.

IMF (International Monetary Fund). 1989. *International Financial Statistics*. Washington, D.C.

Intercare (Integrated Health Care Services, Inc). 1987. *Health Care Financing in the Philippines: A Country Study*. Manila, Philippines.

Jamison, Dean T., and W. Henry Mosley. 1990. "Selecting Disease Control Priorities in Developing Countries." *The World Bank Health Sector Priorities Review*, Paper HSPR-13. Washington, D.C.

Jimenez, Emmanuel. 1987. *Pricing Policy in the Social Sectors*. Baltimore: Johns Hopkins University Press.

——. 1989. "Social Sector Pricing Revisited: A Survey of Some Recent Controversies." *Proceedings of the World Bank Annual Conference on Development Economics. Supplement to the World Bank Economic Review*. Washington, D.C.

Johnson, Bekki J., and Robert S. Pond. 1988. "AIDS in Africa: A Review of Medical, Public Health, Social Science, and Popular Literature." Summary Report. Report prepared under contract for MISEREOR, Campaign against Hunger and Disease in the World, (Episcopal Organization for Development Cooperation), Aachen, West Germany.

Jolly, Dominique. 1987. "Contribution à l'Etude sur les Problèmes et les Recherches Prioritaires à Entreprendre dans le Cadre de la Gestion et de la Planification des Hôpitaux dans les Pays en Développement." World Bank, Population and Human Resources Department, Population, Health, and Nutrition Division, Washington, D.C.

Jolly, Richard, and Maurice King. 1966. "The Organization of Health Services." In Maurice King, ed., *Medical Care in Developing Countries*. Nairobi: Oxford University Press.

JSI (John Snow, Inc.). 1990. "Papua New Guinea: Health Sector Financing Study Project." Prepared for the Papua New Guinea Department of Health under contract with the Asian Development Bank, TA 1091-PNG. Boston.

King, Maurice, ed. 1966. *Medical Care in Developing Countries*. Nairobi: Oxford University Press.

Kleczkowski, B. M., R. H. Elling, and D. L. Smith. 1984. "Health System Support for Primary Health Care." WHO Public Health Papers 80. Geneva.

KMIC (Korea Medical Insurance Corporation). 1987. *'86 Medical Insurance Statistical Yearbook*. Eighth Edition. Seoul.

Knaus, William A., Elizabeth A. Draper, Douglas P. Wagner, and Jack E. Zimmerman. 1985. "APACHE II: A Severity of Disease Classification System." *Critical Care Medicine* 13(10):818–29.

Kutzin, Joseph D. 1989. "Report of Consultancy in Economic and Financial Analysis, Jamaican Hospital Restoration Project." Contract APO-29562 (RR), Pan American Health Organization. Project HOPE Center for Health Affairs: Chevy Chase, Md.

Laffont, Jean Jacques. 1989. "A Brief Overview of the Economics of Incomplete Markets." *Economic Record* 65(3):54–65.

Lewis, Maureen A. 1987. "A Study of Hospital Fees in the Dominican Republic." Working Paper 3714-01 and Research for Child Health (REACH) Project Report. Urban Institute, Washington, D.C.

——. 1988. "Government Policy and the Effectiveness of User Charges in Jamaican Hospitals." Working Paper 3714-05. Urban Institute, Washington, D.C.

——. Forthcoming. "User Fees in Public Hospitals: Comparison of Three Country Case Studies." *Economic Development and Cultural Change.*

Lewis, Maureen A., and Clover Parker. 1991. "Policy and Implementation of User Fees in Jamaican Public Hospitals." *Health Policy* 18(1):57–85.

Lewis, Maureen A., Margaret B. Sulvetta, and Gerard M. LaForgia. 1990. "Estimating Public Hospital Costs by Measuring Resource: A Dominican Case." Working Paper 3714-06, Urban Institute, Washington, D.C.

McGreevey, William P. 1988. "The High Costs of Health Care in Brazil." *PAHO Bulletin* 22(2):145–66.

——. 1990. *Social Security in Latin America: Issues and Options for the World Bank.* World Bank Discussion Paper 110. Washington, D.C.

Mann, Judith K., and Donald E. Yett. 1968. "The Analysis of Hospital Costs: A Review Article." *Journal of Business* 41(2):191–202.

Marquez, Patricio V. 1989. "Uruguay: Population, Health, and Nutrition Sector Memorandum." World Bank, Latin America and the Caribbean, Country Department IV, Human Resources Sector Operations Division, Washington, D.C.

——. 1990. "Containing Health Costs in the Americas." *Health Policy and Planning* 5(4):299–315.

Meesook, O. A. 1984. *Financing and Equity in the Social Sectors in Indonesia: Some Policy Options.* World Bank Staff Working Paper 703. Washington, D.C.

Mills, Anne. 1985. "Survey and Examples of Economic Evaluation of Health Programs in Developing Countries." *World Health Quarterly* 28(4):402–32.

——. 1987. "The Financing and Economics of Hospitals in Developing Countries: Key Issues and Research Questions." PHN Technical Note 87-20, World Bank, Population and Human Resources Department, Population, Health, and Nutrition Division, Washington, D.C.

——. 1989. Distribution of All Reported Inpatient Cases and Deaths by Disease Category and by Region, Malawi, 1986. Ministry of Health, Lilongwe, Malawi.

——. 1990a. "The Economics of Hospitals in Developing Countries. Part I: Expenditure Patterns." *Health Policy and Planning* 5(2):107–17.

——. 1990b. "The Economics of Hospitals in Developing Countries. Part II: Costs and Sources of Income." *Health Policy and Planning* 5(3):203–18.

——. 1991. "The Costs of the District Hospital: A Case Study from Malawi." PRE Working Paper 742, World Bank, Population and Human Resources Department, Population, Health, and Nutrition Division, Washington, D.C.

Mills, Anne, and Michael Drummond. 1987. "Value for Money in the Health Sector: The Contribution of Primary Health Care." *Health Policy and Planning* 2(2):107–28.

Mills, Anne, Jean Njoloma, and Steve Chisimbi. 1989. "The Role of the Hospital in the District Health Sector: The Pattern of Resource Allocation and Unit Costs

in a Sample of Malawian Districts." World Bank, Population and Human Resources Department, Population, Health, and Nutrition Division, Washington, D.C.

Milner, N. P. H. 1985. "Logistics in the Role of Hospitals at the First Referral Level." Working Paper SHS/EC/85/WP/II.E.1 from the WHO Expert Committee on the Role of Hospitals at the First Referral Level. Geneva.

Mobarak, Almotaz Billah, Mostafa T. Hammany, Osman M. El-Zehmaity, and M. Nabil Nassar. 1981. Diarrheal Disease Control Study. Strengthening Rural Health Delivery Project Final Report on Phase I to USAID and WHO. Cairo.

Mohr, Penny. 1986. "Recommendations for User Fees for Medical Services in Grenada." Memorandum 2, Project HOPE. St. George's, Grenada.

MPHSA (Ministry of Public Health and Social Affairs). 1987a. *Rapport Annuel sur l'Autonomie de Gestion des Hôpitaux Nationaux du Mali.* Koulouba, Mali.

——. 1987b. Rapport de Gestion, Hôpital National du Point "G." Bamako, Mali.

Mwabu, Germano M. 1989. "Referral Systems and Health Care Seeking Behavior of Patients: An Economic Setting." *World Development* 17(1):85–92.

Mwabu, Germano M., and Wilfred M. Mwangi. 1986. "Health Care Financing in Kenya: A Simulation of Welfare Effects of User Fees." *Social Science and Medicine* 22(7):763–67.

NCHS (National Center for Health Statistics). 1989. *Health, United States, 1988.* Annual published by the United States Department of Health and Human Services, Public Health Service, Office of Health Research, Statistics, and Technology, National Center for Health Statistics, National Center for Health Services Research. Rockville, Md.

Newbrander, William C. 1987. "Papua New Guinea's Expenditure on Hospitals: Policy and Practice Since Independence." *Health Policy and Planning* 2(3):227–35.

Newhouse, Joseph P. 1970. "Toward a Theory of Nonprofit Institutions: An Economic Model of a Hospital." *American Economic Review* 60(1):64–74.

OECD (Organization for Economic Cooperation and Development). 1985. *Measuring Health Care 1960–1980: Expenditures, Costs and Performance.* OECD Social Policy Studies 2. Paris.

——. 1987. *Financing and Delivering Health Care: A Comparative Analysis of OECD Countries.* OECD Social Policy Studies 4. Paris.

——. 1990. *Health Care Systems in Transition: The Search for Efficiency.* OECD Social Policy Studies 7. Paris.

Over, Mead. 1981. "On the Care and Feeding of a Gift Horse: The Recurrent Cost Problem and Optimal Reduction of Recurrent Inputs." Research Memorandum Series 79. Williams College, Center for Development Economics. Williamstown, Mass.

Over, Mead, S. Bertozzi, J. Chin, B. N'Galy, and K. Nyamuryekung'e. 1988. "The Direct and Indirect Cost of HIV Infection in Developing Countries: The Cases of Zaire and Tanzania." In Alan F. Fleming, Manuel Carballo, David W. FitzSimons, Michael R. Bailey, and Jonathan Mann, eds., *The Global Impact of AIDS.* New York: Alan R. Liss.

Over, Mead, Randall P. Ellis, Joyce H. Huber, and Orville Solon. 1992. The Consequences of Adult Ill Health." In Richard G. A. Feacham, Tord Kjellstrom, Christopher J. L. Murray, Mead Over, and Margaret A. Phillips, eds., *The Health of Adults in the Developing World.* New York: Oxford University Press.

Overholt, Catherine. 1987. "User Fees in Honduran Hospitals and Health Centers: Policy and Experience." Research for Child Health (REACH) Project Report. John Snow, Inc., Arlington, Va.

Overholt, Catherine, Hopkins Holmberg, Colette Marcellin, and Lazarus Murahwa. 1990. "Improving the Cost Recovery System: Assessment of Fee Structure and Billings and Collections Procedures." Prepared for the Zimbabwe Ministry of Health under contract with the World Bank, John Snow, Inc., Boston.

Pabón Lasso, Hipolito. 1986. "Evaluating Hospital Performance through Simultaneous Application of Several Indicators." *Bulletin of the Pan American Health Organization* 20(4):341–57.

Paine, L. H. W., and F. Siem Tjam. 1988. *Hospitals and the Health Care Revolution.* Geneva: World Health Organization.

Palmer, P. E. S. 1985. "The World Health Organization: Basic Radiological System." *Radiology* 51(597):169–78.

Pande, B. R. 1987. "Health Care Financing Country Paper: Nepal." Country Paper prepared for the Regional Seminar on Health Care Financing, Asian Development Bank, Manila, Philippines.

Pangu, K. A., L. de Buysscher, R. Vervoort, and M. Réveillon. 1987. "Etude de Fonctionnement et de Gestion des Hôpitaux du Bénin." Medicus Mundi, Belgium.

Parkin, D. M., and S. M. Moss. 1986. "An Evaluation of Screening Policies for Cervical Cancer in England and Wales Using a Computer Simulation Model." *Journal of Epidemiological and Community Health* 40(2):143–53.

Pauly, Mark V. 1987. "Nonprofit Firms in Medical Markets." *AEA Papers and Proceedings* 77(2):257–74.

People's Republic of China, Ministry of Health. 1986. "Survey of Urban Health Services in Hubei, Sichuan, and Jilin Provinces." Beijing.

Pitt, Mark M., and Mark R. Rosenzweig. 1990. "Estimating the Intrahousehold Incidence of Illness: Child Health and Gender-Inequality in the Allocation of Time." *International Economic Review* 31(4):969–90.

Prescott, Nicholas, ed. 1991. *Indonesia: Health Planning and Budgeting.* A World Bank Country Study. Washington, D.C.

Prescott, Nicholas, and Dean T. Jamison. 1985. "The Distribution and Impact of Health Resource Availability in China." *International Journal of Health Planning and Management* 1(1):45–56.

PRIDES (Programa de Investigación y Desarrollo en Sistemas de Salud). 1980. *Evaluación de la Gestion Financiera y Beneficio Social de los Hospitales de Referencia en Colombia: 1975–1980.* Cali, Colombia.

Prost, André, and Nicholas Prescott. 1984. "Cost-Effectiveness of Blindness Prevention by the Onchocerciasis Control Programme in Upper Volta." *Bulletin of the World Health Organization* 62(5):795–802.

Pust, Ronald E. 1985. "Analyzing Referral Patterns to Rural Provincial Hospitals." *Tropical Doctor* 15(2):95–97.

Quick, Jonathan D., ed. 1982. *Managing Drug Supply*. Boston: Management Sciences for Health.

Ramsey, Frank P. 1927. "A Contribution to the Theory of Taxation." *Economic Journal* 37(1):47–61.

Rao, Bhasker. 1990. "Perinatal Mortality." In Helen M. Wallace and Kanti Giri, eds. *Health Care of Women and Children in Developing Countries*. Oakland, Calif.

Raymond, Susan Ueber, Barbara Lewis, Paul Meissner, and Jeremiah Norris. 1987. "Financing and Costs of Health Services in Belize." Health Care Financing in Latin America and the Caribbean Research Report 2. USAID Contract LAC 0632-C-00-5137-00. Stony Brook, N.Y.

Robertson, Robert L. 1985. "Review of Literature on Costs of Health Services in Developing Countries." PHN Technical Note 85-21. World Bank, Population and Human Resources Department, Population, Health, and Nutrition Division, Washington, D.C.

Rodrigues, José. 1989a. "What is Happening to Hospital Utilization in Brazil?" *Health Policy and Planning* 4(4):354–56.

——. 1989b. "Hospital Utilization and Reimbursement Method in Brazil." *International Journal of Health Planning and Management* 4(1):3–15.

Rosenzweig, Mark R., and T. Paul Schultz. 1983. "Estimating a Household Production Function: Heterogeneity, the Demand for Health Inputs, and Their Effects on Birth Weight." *Journal of Political Economy* 91(5):723–46.

Russell, Sharon S., Gretchen Gwynne, and Michael Trisolini. 1988. "Health Care Financing in St. Lucia and Costs of Victoria Hospital." Health Care Financing in Latin America and the Caribbean Research Report 5. USAID Contract LAC 0632-C-00-5137-00. Stony Brook, N.Y.

Sabatier, Renée, and Jon Tinker. 1989. *AIDS and the Third World*. Panos Institute, London.

Saxenian, Helen. Forthcoming. *Brazil: Women's Reproductive Health*. A World Bank Country Study. Washington, D.C.

Selowsky, Marcelo. 1979. *Who Benefits from Government Expenditure? A Case Study from Colombia*. Washington, D.C.: World Bank.

Seth, Mira, and Om Prakash Gupta. 1987. "Country Paper on Health Care Financing: India." Country Paper prepared for the Regional Seminar on Health Care Financing, Asian Development Bank. Manila, Philippines.

Sewankambo, Nelson K. 1989. "Care of HIV-infected People in the Developing World: Practical Aspects." *AIDS* 3 (Supp. 1):S195–99.

Shepard, Donald S. 1988. "Analysis and Recommendations on Health Financing in Rwanda." Report prepared for the World Bank, Washington, D.C., and *Projet Santé Familiale*, Rwanda.

Shepard, Donald S., Logan E. Brenzel, and Karen T. Nemeth. 1986. "Cost-Effectiveness of Oral Rehydration Therapy for Diarrheal Disease." PHN Technical Note 86-26. World Bank, Population and Human Resources Department, Population, Health, and Nutrition Division, Washington, D.C.

Shepard, Donald S., Layes Sanoh, and Emmou Coffi. 1986. "Cost Effectiveness of the Expanded Program on Immunization in the Ivory Coast: A Preliminary Assessment." *Social Science and Medicine* 22(3):369–77.

Shepard, Donald S., Taryn Vian, and Eckhard F. Kleinau. 1990. "Health Insurance in Zaire." PRE Working Paper 489. World Bank, Africa Technical Department, Population Health, and Nutrition Division, Washington, D.C.

Shepard, Donald S., Julia Walsh, Wolfgang Munar, Laura Rose, Rodrigo Guerrero, Luis F. Cruz, Guillermo Reyes, Gail Price, and Carlos Solarte. 1991. "Cost-Effectiveness of Ambulatory Surgery in Cali, Colombia." Working Paper 1, Department of Population and International Health. Harvard School of Public Health, Boston.

Sherman, Roger. 1989. *The Regulation of Monopoly*. New York: Cambridge University Press.

Shimmura, Kazuya. 1988. "Effects of Different Remuneration Methods on General Medical Practice: A Comparison of Capitation and Fee-for-Service Payment." *International Journal of Health Planning and Management* 3(4):245–58.

Starrs, Ann, ed. 1987. *Preventing the Tragedy of Maternal Deaths*. Nairobi: International Safe Motherhood Conference.

Supit, B. A. 1985. "The First Referral Level Hospital Support for the Local Health System Infrastructure (The Bethesda Tomohon Hospital Involvement and Experience, North Sulawesi, Indonesia)." Working Paper SHS/EC/85/WP/II.D.1 from the WHO Expert Committee on the Role of Hospitals at the First Referral Level. Geneva.

Swerdlow, M., and Jan Stjernswärd. 1982. "Cancer Pain Relief: An Urgent Problem." *World Health Forum* 3(3):325–30.

Thomason, J. A., D. Makavo, K. L. Loh, and R-L. Kolehmainen-Aitken. 1987. "Health Care Financing: Papua New Guinea." Country Paper prepared for the Regional Seminar on Health Care Financing, Asian Development Bank. Manila, Philippines.

Torrance, George. 1986. "Measurement of Health State Utilities for Economic Appraisal." *Journal of Health Economics* 5(1):1–30.

United Medical Enterprises. 1988. "Lesotho Health Financing Study." Report prepared for the Ministry of Health, Kingdom of Lesotho. London.

Vélez Gil, Adolfo, and Graciela Pardo de Vélez. 1985. *Investigación de Modelos de Atención en Cirugía*. Cali, Colombia.

Vita, Michael G. 1990. "Exploring Hospital Production Relationships with Flexible Functional Forms." *Journal of Health Economics* 9(1):1–21.

Vogel, Ronald J. 1988. *Cost Recovery in the Health Care Sector: Selected Country Studies in West Africa*. World Bank Technical Paper 82. Washington, D.C.

Waddington, Catriona J., and K. A. Enyimayew. 1989. "A Price to Pay: The Impact of User Charges in Ashanti-Akim District, Ghana." *International Journal of Health Planning and Management* 4(1):17–47.

——. 1990. "A Price to Pay, Part 2: The Impact of User Charges in the Volta Region of Ghana." *International Journal of Health Planning and Management* 5(4):287–312.

Wagstaff, Adam. 1989a. "Econometric Studies in Health Economics: A Survey of the British Literature." *Journal of Health Economics* 8(1):1–51.

——. 1989b. "Estimating Efficiency in the Hospital Sector: A Comparison of Three Statistical Cost Frontier Models." *Applied Economics* 21(5):659–72.

Walker, Godfrey J. A. 1986. "Hospitals and Primary Health Care." World Bank, Population and Human Resources Department, Population, Health, and Nutrition Division, Washington, D.C.

Watters, D. A. K., I. H. Wilson, J. R. Sinclair, and N. Ngandu. 1989. "A Clinical Sickness Score for the Critically Ill in Central Africa." *Intensive Care Medicine* 15:467–70.

Weaver, Marcia, Kadi Handou, and Zeynabou Mohamed. 1990a. "Study of the Amount of Costs Recovered in the National Hospitals and CHDs of Niger." Prepared under USAID Project 683-0254. Abt Associates. Bethesda, Md.

——. 1990b. "Patient Surveys at Niamey National Hospital: Results and Implications for Reform of Hospital Fees." Prepared under USAID Project 683-0254. Abt Associates. Bethesda, Md.

Wellstart/The San Diego Lactation Program. 1990. "The Lactation Management Education Program for the Philippines: Its Contribution to National Breastfeeding Promotion." San Diego.

WHO (World Health Organization). 1979. "Laboratory Services at Primary Health Care Level." WHO/LAB/79.1. Geneva.

——. 1985. *The Selection of Essential Drugs*. Technical Report Series 722. Geneva.

——. 1986. "Key Issue 6: Imaging Technologies." Global Programme for Appropriate Health Care Technology. Danish Hospital Institute, Copenhagen, Denmark.

——. 1987a. *Hospitals and Health for All*. Report of a WHO Expert Committee on the Role of Hospitals at the First Referral Level. Technical Report Series 744. Geneva.

——. 1987b. "Interregional Meeting on the Maintenance and Repair of Health Care Equipment." WHO/SHS/NHP/87.5. Nicosia, Cyprus.

——. 1987c. "Laboratory Services at the Primary Health Care Level." WHO/LAB/87.2. Geneva.

WHO/GPA (World Health Organization, Global Programme on AIDS). 1992. "The Current and Future Dimensions of the HIV/AIDS Pandemic." WHO/GPA/RES/SFI/92.1. Geneva.

Williams, Alan. 1985. "Economics of Coronary Artery Bypass Grafting." *British Medical Journal* 291 (August 3):326–29.

Wong, Holly. 1989. "Cost Analysis of Niamey Hospital." Prepared under USAID Project 683-0254. Abt Associates. Bethesda, Md.

World Bank. 1984. *Zambia: Population, Health and Nutrition Sector Review*. Report 4715-ZA. Southern Africa Department, Population and Human Resources Operations Division, Washington, D.C.

———. 1985. *Somalia: Population, Health and Nutrition Sector Review.* Report 5468-SO. Eastern Africa Department, Population and Human Resources Operations Division, Washington, D.C.

———. 1988. *Brazil: Public Spending on Social Programs; Issues and Options.* Report 7086-BR. Latin America and the Caribbean Country Department I, Population and Human Resources Operations Division, Washington, D.C.

———. 1989a. *Korea: Health Insurance and the Health Sector.* Report 7412-KO. East Asia and Pacific Country Department I, Population and Human Resources Operations Division, Washington, D.C.

———. 1989b. *Nepal: Social Sector Strategy Review.* Report 7498-NEP. South Asia Country Department I, Population and Human Resources Operations Division, Washington, D.C.

———. 1990. *Issues and Options in Health Financing in Turkey.* Report 8042-TU. Europe and Central Asia Country Department I, Human Resources Operations Division, Washington, D.C.

———. 1991. *Kenya: Human Resources: Improving Quality and Access.* Report 9023-KE. Eastern Africa Department, Population and Human Resources Operations Division, Washington, D.C.

Wouters, Annemarie. Forthcoming. "The Cost and Efficiency of Public and Private Health Care Facilities in Ogun State, Nigeria. *Health Economics.*

Yang, Bong-min. 1991. "Health Insurance in Korea: Opportunities and Challenges." *Health Policy and Planning* 6(2):119–29.

Young, Mary. 1989. "Alternatives to Hospital Care for Chronic Illnesses." World Bank, Population and Health Resources Department, Population, Health, and Nutrition Division, Washington, D.C.

Yu, Seung-Hum. 1983. "Study on Hospital Care Services between Insured and Non-insured Patients for Selected Diagnoses in Korea." *Yonsei Medical Journal* 24(1):6–32.

Zimmerman, Jack E., ed. 1989. "APACHE III Study Design: Analytic Plan for Evaluation of Severity and Outcome." *Critical Care Medicine* 17(12):S169–221.

Index

The World Bank

Hospitals in developing countries absorb more resources than any other kind of recurrent government spending on health. But although hospitals are essential, other types of intervention are more cost-effective in providing some of the most needed services. In many countries, much health expenditure is wasted through misallocation and by technical and managerial inefficiency within hospitals. Improvements in the use of resources by hospitals could yield tremendous benefits. This book advocates drawing hospitals in developing countries into the policy dialogue on the use of public resources in the health sector.

The authors focus on the broad issues of resource allocation, costs, and financing. They advocate using a combination of hospital cost data with service statistics to highlight areas in which waste can be reduced. Hospital costs and efficiency are examined through accounting based studies and analysis of statistical cost functions. The authors recommend principles for pricing hospital services and alternatives for hospital financing through user charges and health insurance. These recommendations draw on both economic theory and a review of country experience through case studies and summaries of country case examples.

Finally, the book recommends coordination between hospital and nonhospital forms of treatment. It suggests that some hospital procedures could be equally well performed in other settings but at less cost.

Howard Barnum is a senior health economist and **Joseph Kutzin** is a consultant health economist, both in the Population, Health, and Nutrition department at the World Bank.

The Johns Hopkins University Press

Cover design by Beni Chibber-Rao

ISBN 0-8018-4532-7