THE IMPACT OF RESOURCE ALLOCATION AND PURCHASING REFORMS ON EQUITY

Paolo Carlo Belli

September 2004
Health, Nutrition and Population (HNP) Discussion Paper

This series is produced by the Health, Nutrition, and Population Family (HNP) of the World Bank's Human Development Network (HNP Discussion Paper). The papers in this series aim to provide a vehicle for publishing preliminary and unpolished results on HNP topics to encourage discussion and debate. The findings, interpretations, and conclusions expressed in this paper are entirely those of the author(s) and should not be attributed in any manner to the World Bank, to its affiliated organizations or to members of its Board of Executive Directors or the countries they represent. Citation and the use of material presented in this series should take into account this provisional character. For free copies of papers in this series please contact the individual authors whose name appears on the paper.

Enquiries about the series and submissions should be made directly to the Editor in Chief. Submissions should have been previously reviewed and cleared by the sponsoring department which will bear the cost of publication. No additional reviews will be undertaken after submission. The sponsoring department and authors bear full responsibility for the quality of the technical contents and presentation of material in the series.

Since the material will be published as presented, authors should submit an electronic copy in a predefined format as well as three camera-ready hard copies (copied front to back exactly as the author would like the final publication to appear). Rough drafts that do not meet minimum presentational standards may be returned to authors for more work before being accepted.

The Editor in Chief of the series is Alexander S. Preker (apreker@worldbank.org); For information regarding this and other World Bank publications, please contact the HNP Advisory Services (healthpop@worldbank.org) at: Tel (202) 473-2256; and Fax (202) 522-3234.
Abstract: This paper introduces a conceptual framework to investigate the equity consequences of resource allocation, strategic purchasing and payment system reforms (RAP) in health, and reviews the empirical literature that has analyzed the impact of these reforms on equity.

In the first part, the paper reviews the existing evidence on the distribution of health and health care utilization across socio-economic groups in developing countries. Such evidence shows a striking consistency in the association between poverty and poor health, and that the distribution of health benefits across socio-economic groups is highly unequal, favoring the wealthier segments of the population. At the same time it shows that the issue of quality of services is as important as the issue of their accessibility.

Then the paper addresses the meaning of equity in health care: we argue that the concept of equity and that of an equitable distribution of resources in health are not as straightforward as they appear.

In the second part, the paper reviews the core elements of RAP reforms, and it presents a survey of the empirical literature that has investigated the impact of RAP reforms from an equity perspective. We argue that some RAP reform components, such as the new resource allocation mechanisms, bear a strong pro-poor potential, but that other components can lead to the emergence of trade-offs between conflicting objectives, such as efficiency and equity. For instance, the new purchasing contracts and payment systems devised to enhance efficiency can create adverse equity effects. Conceptually, these trade-offs are significant and operating at several levels.

Keywords: resource allocation and purchasing, health care financing, equity, health reforms, resource allocation, payment systems

Disclaimer: The findings, interpretations and conclusions expressed in the paper are entirely those of the authors, and do not represent the views of the World Bank, its Executive Directors, or the countries they represent.

Correspondence Details: Paolo Carlo Belli; IHSD 27 Old Street, London, EC1V 9HL; Tel: (++44) 78-81994115; Email: pbelli@hsph.harvard.edu, or Paolo.belli@ihsd.org
Table of Contents

FOREWORD.................................................................................................................................VII

ACKNOWLEDGEMENTS ............................................................................................................... IX

INTRODUCTION........................................................................................................................... 1

EVIDENCE ON EXISTING INEQUALITIES IN HEALTH, AND IN HEALTH BENEFITS FROM PUBLIC SERVICES IN DEVELOPING COUNTRIES .............. 2

THE DISTRIBUTION OF BENEFITS FROM PUBLIC EXPENDITURE ON HEALTH. EVIDENCE FROM DEVELOPING COUNTRIES. ................................................................. 7

WHAT IS THE MEANING OF AN EQUITABLE DISTRIBUTION OF HEALTH AND HEALTH BENEFITS? ....................................................................................................................... 12

EQUITY IN HEALTH CARE. THE DIFFERENT PRINCIPLES IN THE LITERATURE ......................... 16

WHAT IS THE MEANING OF RAP REFORMS? .............................................................................. 20

MOTIVATION OF RAP REFORMS ............................................................................................... 25

THE RELATIONSHIP BETWEEN RAP REFORMS AND PRIVATIZATION ........................................ 29

RAP REFORM COMPONENTS: A REVIEW OF THE EVIDENCE OF THEIR IMPACT ON EQUITY ........ 29

Reforms of the Resource Allocation Criteria ........................................................................ 30

The Priority Setting Component of RAP Reforms ................................................................ 34

Purchasing/contracting reforms ............................................................................................ 36

Payment systems’ reform ........................................................................................................ 40

The Use of RAP Arrangements in Specific Pro-poor Health Interventions .............................. 44

Supply-side incentives .............................................................................................................. 44

Demand side incentives ........................................................................................................... 46

TRADE-OFFS IN THE PURCHASING AND PAYMENT SYSTEM COMPONENTS OF RAP REFORMS ................................................................................................................................. 47

PURCHASERS: CONSUMER CHOICE AND CONSUMER-RESPONSIVENESS VERSUS NO RISK SELECTION .................................................................................................................. 49

EFFICIENCY VERSUS THE RIGHT TO EQUAL TREATMENT ...................................................... 52

COST CONSCIOUSNESS VERSUS QUALITY ENHANCEMENT AND NO RISK SELECTION .......... 54

CONCLUSION ................................................................................................................................ 55

ANNEX 1 UNDERSTANDING THE LINK BETWEEN SOCIO-ECONOMIC DEPRIVATION AND HEALTH. EVIDENCE FROM INDUSTRIALIZED COUNTRIES ............................................................................. 57

PHYSIOLOGICAL ........................................................................................................................ 57

BEHAVIORAL ................................................................................................................................ 58

PSYCHOLOGICAL ........................................................................................................................ 58

HEALTH INEQUALITIES AND THEIR LINK WITH HEALTH SYSTEMS’ CHARACTERISTICS ............... 58

ANNEX 2 EQUITY IN HEALTH: THE HUMAN RIGHTS’ PERSPECTIVE ........................................ 61
Great progress has been made in recent years in securing better access and financial protection against the cost of illness through collective financing of health care. This publication – *The Impact of Resource Allocation and Purchasing Reforms on Equity* by Paolo Carlo Belli – is part of a series of Discussions Papers that review ways to make public spending on health care more efficient and equitable in developing countries through strategic purchasing and contracting services from nongovernmental providers.

Promoting health and confronting disease challenges requires action across a range of activities in the health system. This includes improvements in the policymaking and stewardship role of governments, better access to human resources, drugs, medical equipment, and consumables, and a greater engagement of both public and private providers of services.

Managing scarce resources and health care effectively and efficiently is an important part of this story. Experience has shown that, without strategic policies and focused spending mechanisms, the poor and other ordinary people are likely to get left out. The use of purchasing as a tool to enhance public sector performance is well documented in other sectors of the economy. Extension of this experience to the health sector is more recent and lessons learned are now being successfully applied to developing countries.

The shift from hiring staff in the public sector and producing services “in house” from non governmental providers has been at the center of a lively debate on collective financing of health care during recent years. Its underlying premise is that it is necessary to separate the functions of financing health services from the production process of service delivery to improve public sector accountability and performance.

In this Discussion Paper, Belli examines in greater detail the equity dimensions of the new resource allocation, purchasing and payment systems. Two key observations which emerge from his review include: (a) the significant existing differentials in health expenditure per capita between richer and poorer countries; and (b) the pro rich utilization of health services that is observed in most countries, despite the fact that the poor bear a disproportionate morbidity and mortality. More pro poor resource allocation and purchasing of health services is needed to address this problem.

*Alexander S. Preker*

Lead Economist
Editor of HNP Publications
ACKNOWLEDGEMENTS

My warmest thanks go to all of you who have inspired, assisted, commented, challenged and critiqued my work. This paper is based upon a number of studies which several authors have contributed to over the years. I would like to thank all these and colleagues and others who shared with me their work, and specifically Cristian Baeza, Peter Berman, Thomas Bossert, Josep Figueras, Michele Gragnolati, David Gwatkin, John Langenbrunner, Benjamin Loevinsohn, Alex Preker, Susan Stout, Adam Wagstaff, Diana Weil, and Abdo Yazbeck. I would like to also thank all the other authors, whom I did not have an opportunity to meet and whose work I am nonetheless referring to in this paper. The interpretations and conclusions expressed in this paper are exclusively mine.

The author of this Report is grateful to the World Bank for having published it as an HNP Discussion Paper.
INTRODUCTION

This paper introduces a conceptual framework to investigate the equity consequences of resource allocation, strategic purchasing and payment system (RAP) reforms in health, and it reviews the existing evidence on the impact of these reforms.

In the first part, the paper reviews the existing evidence on the distribution of health and health care utilization across socio-economic groups in developing countries. Such evidence shows a striking consistency in the association between poverty and poor health, and that the distribution of health benefits across socio-economic groups is highly unequal, favoring the wealthier segments of the population. Then, the paper addresses the meaning of equity in health: we argue that the concept of an equitable distribution of health resources is not as straightforward as it first appears. Hence, assessing changes at the margin may be difficult, and different equity principles and measures would lead to opposite conclusions.

In the second part, the paper reviews the core elements of RAP reforms, and it presents a survey of the empirical literature that has investigated the impact of RAP reforms from an equity perspective. We argue that some RAP reform components, such as the new resource allocation mechanisms, bear a strong pro-poor potential, but that other components, such as the new contracting mechanisms and payment systems for providers can lead to the emergence of new trade-offs. Trade-offs exist among potentially equally desirable aims in the sense that each reform component is appropriate to achieve only one of those aims, and, by doing so, in general, it moves the health system further away from achieving other equally desirable aims. For instance, the new purchasing contracts and payment systems do enhance efficiency, but can create adverse equity effects. Conceptually, these trade-offs are significant and operating at several levels.

Note that the health policy literature does not, on the whole, talk in the language of trade-offs. This has the (probably unintended) implication that it is possible to achieve all objectives simultaneously, if only the right policy intervention, being it a specific resource mobilization or resource allocation mechanism, or a specific payment system for providers, can be found. In contrast, we argue that in general each health reform component comes with potential negative “side” effects. Given the existence of trade-offs, health reforms should contemplate a plurality of interventions (in resource mobilization and allocation, payment system, management, etc.), able to reinforce each other and at the same time to correct the potential negative impact of each intervention taken in isolation.

One of the main conclusions of the paper is that the evidence available on the RAP reforms’ impact on equity is still largely inconclusive. The lack of a scientifically rigorous evidence base to evaluate the RAP reforms is partially due to intrinsic measurement difficulties. Aggregate measures of equity, especially those concerning distribution of health outcomes, change slowly, while most RAP reforms in developing countries are recent. The level and distribution of education or income per capita and several other variables are perhaps more important than utilization of health services for explaining health differentials across socioeconomic groups. But because they are also continuously in evolution, it is extremely difficult to ascribe to health
reforms some of the distributional changes observed. However, the lack of conclusive evidence is also caused by an insufficient attention to monitoring and evaluating the reforms, and specifically to the equity dimension of the reforms. For instance, one can hardly find any data to evaluate the impact of RAP reforms where the change in utilization of essential health services is measured disaggregated by socio-economic groups, with an explicit focus on the benefits for the poor. We hope that in the future new and more conclusive evidence will soon be brought to bear on some of the conceptual arguments and hypotheses presented here. In the meantime, by necessity many conclusions of our review are still based on conceptual arguments, rather than on a solid empirical basis. From theory we know how different health policy interventions are likely to change the incentive structure within the health system, and so what impact they can be expected to produce. From that knowledge we can develop hypotheses to be tested in future empirical research.

**EVIDENCE ON EXISTING INEQUALITIES IN HEALTH, AND IN HEALTH BENEFITS FROM PUBLIC SERVICES IN DEVELOPING COUNTRIES**

Most of the evidence available on distribution of health and health resources across socio-economic groups in developing countries is presented in studies not related to specific health reforms, or to evaluation of RAP reforms. We briefly present such evidence in this section.

Regardless of the diversity of the existing studies in terms of measurement approaches to equity and poverty, study design and geographical focus, the empirical evidence indicates a striking consistency in the association between poverty and poor health. Beaglehole and Bonita (1997, p.1) assert that poverty is the most important cause of preventable death, disease and disability.

Inter country comparisons present evidence on health inequalities across countries, and their evolution over time. Some of the key findings are the following:

- The bulk of mortality and morbidity affects people living in the developing world. 98% of the deaths between birth and 15 years and 83% of the deaths between 15 and 59 years occur in the developing world (Murray and Lopez, 1996).

- These inequalities have not decreased over time. Gwatkin, 2000b, analyzing World Bank data set, concludes that in the period 1970-1990 infant mortality differentials across countries declined in absolute terms, but that they increased in relative terms ("While infant mortality in 1970 was around 6.5 times as high in the poorest as in the richest countries, it was over 11 times as high by 1990 (Gwatkin, 2000b, p. 19)"). He also points at the fact that, unlike under-5 mortality figures, adult mortality differentials did unequivocally narrow over time.

- Over 30% (53% in sub-Saharan Africa) of deaths in the developing world occur in children younger than 5 years. In terms of DALYs lost the three main groups of disease are perinatal disorders, lower respiratory infections, and diarrhoeal diseases, all of which are prevalent in the developing world, especially in the poorest countries and among the younger and poorer segments of the population (see discussion hereafter).
Life expectancy is strongly associated with income per capita. However, the relationship is not linear (Preston, 1980) and different countries with similar level of income achieve widely different results in mortality rates and life expectancies. Among the good performers relative to their level of socio-economic development, we find Cuba, Sri Lanka, the Indian State of Kerala, and Costa Rica. For instance in 1997 Cuba and Costa Rica had a per capita Gross Domestic Product respectively equal to $3,100 and $3,810, an under-five mortality equal to 9 and 13 respectively, and a life expectancy equal to 75.7, 76.0 years, which are indicators comparable to those of the wealthiest countries of West Europe and North America (World Bank, 1999). In 1998 life expectancy in Sri Lanka was 73 years, and the infant mortality rate was 15.4 per 1,000 live births; in 2001 income (GDP) per capita was only US$837. By comparison, in Brazil GDP per capita was $2,940, but 20 percent of children were still reported as undernourished, the infant mortality rate was 50 per 1,000 live births, and life expectancy was about 67 years (World Bank 1999).

More recent literature has looked into the relationship between poverty and health within countries. The main findings of this literature are three. A) Mortality is related as much to socio-economic inequalities (in asset ownership, expenditure, income, education attainment, etc.) within countries, as it is to differences in absolute wealth/income amongst them. B) National mortality rates are positively correlated with levels of deprivation (for example, Anand and Ravallion, 1993, show that a significant part of the impact of income per capita on mortality withers away once we explicitly consider the share of the population in poverty as an explanatory variable). C) There is increasing evidence that demonstrates an independent positive impact of social capital on health outcomes (see Annex 1, and the evidence presented by Subramanian et. al., 2001; Kawachi et. al., 1999).

Investigations of health and health utilization differentials across socio-economic groups in developing countries are based on two major sources of evidence: Living Standard Measurement Surveys (LSMS) and Demographic and Health Surveys (DHS). While LSMS contain very accurate information of households’ level of consumption and expenditure, there is less detailed information on health seeking behavior and health outcomes. By contrast, DHS data contain very accurate information over health status and health service utilization, especially for reproductive, maternal and child care services, but no information on households’ income or expenditure. The two above main sources of evidence are discussed respectively by Gwatkin et al. (2000) and by Wagstaff (2000).

On the basis of DHS data collected in over 40 developing countries, Gwatkin et al. (2000) analyze inequalities across socio-economic groups in a) infant and under 5 mortality; b) levels of malnutrition; c) incidence of diarrhea and acute respiratory infection; d) fertility rates; e) information on HIV/AIDS; and f) use of Maternal and Child Health services. They adopt a wealth index developed by Filmer and Pritchett (1998) to divide the population into different groups.
wealth quintiles according to reported asset ownership, and show that disparities between poor and non-poor vary enormously across countries. The tables below present some of their key findings. On average across countries, a child born in a household belonging to the lowest wealth quintile is roughly twice as likely to die before reaching 5 years than a child born in a household from the highest wealth quintile. Inequalities are highest in the Latin American/Caribbean region.

### Table 1 and 2: Intra-Country Disparities in Infant and Under-Five Mortality

#### 1. Infant Mortality

<table>
<thead>
<tr>
<th>Region</th>
<th>No. of Countries</th>
<th>Poor-Rich Ratio</th>
<th>Concentration Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sub-Saharan Africa</td>
<td>21</td>
<td>Mean 1.67</td>
<td>Range 1.11 to 2.46</td>
</tr>
<tr>
<td>Asia/Near East/North Africa</td>
<td>9</td>
<td>Mean 2.33</td>
<td>Range 1.42 to 3.93</td>
</tr>
<tr>
<td>Latin America/Caribbean</td>
<td>11</td>
<td>Mean 2.66</td>
<td>Range 1.26 to 4.18</td>
</tr>
<tr>
<td>Total</td>
<td>40</td>
<td>Mean 1.87</td>
<td>Range 1.11 to 4.18</td>
</tr>
</tbody>
</table>

#### 2 Under-Five Mortality

<table>
<thead>
<tr>
<th>Region</th>
<th>No. of Countries</th>
<th>Poor-Rich Ratio</th>
<th>Concentration Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sub-Saharan Africa</td>
<td>21</td>
<td>Mean 1.79</td>
<td>Range 1.27 to 2.60</td>
</tr>
<tr>
<td>Asia/Near East/North Africa</td>
<td>9</td>
<td>Mean 2.69</td>
<td>Range 1.69 to 4.60</td>
</tr>
<tr>
<td>Latin America/Caribbean</td>
<td>11</td>
<td>Mean 2.99</td>
<td>Range 1.55 to 4.67</td>
</tr>
<tr>
<td>Total</td>
<td>40</td>
<td>Mean 2.06</td>
<td>Range 1.27 to 4.67</td>
</tr>
</tbody>
</table>

Source: Gwatkin, 2000b

An interesting finding of the study is that countries with lower mortality and morbidity rates among children are in general also characterized by wider disparities across socio-economic groups. This finding seems at odds with the non-linear relationship between income and health, characterized by diminishing returns to scale, and could be explained by one or a combination of the following reasons:

2 Two patterns of the data worth noting are that under-five mortality rates are more concentrated than infant mortality rates (this could result from problems of underreporting of infant mortality), and that the standard deviations are extremely large (so that parameters are not statistically significant for several countries).

3 The concentration index is equal to: \[1 - 2 \int_0^1 MR_i(w_i)dw,\] where MR is the cumulative proportion of mortality rates among children graphed against the cumulative proportion of their households’ wealth (i=1, …, 5). A negative (positive) value of the concentration index indicates inequality favoring the rich (poor).
a) Rates of diffusion of health-promoting innovations are different across different segments of society, first benefiting the wealthier segments of society.

b) Increase in per capita income is associated with increases in income disparities.

c) Societies experience varying degrees of health disparity because some of them put in place policies that make the health-income relationship more elastic while others do not.

d) A potential negative health externality occurs associated with being poor in a richer country.

Using LSMS data, Wagstaff (2000b) compares data on infant and under-5 mortality rates from 9 developing countries. His study finds significant inequalities in infant and under-five mortality rates across quintile-expenditure groups in all countries, although the estimates are not as pronounced as in Gwatkin et al. (2000). For six countries Wagstaff's estimates are directly comparable with Gwatkin’s.4

Table 3: Concentration Indices for Infant and Under-Five Mortality within Six Developing Countries as Measured by Different Studies

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Côte d'Ivoire</td>
<td>-0.107</td>
<td>-0.095</td>
<td>-0.115</td>
<td>-0.096</td>
<td></td>
</tr>
<tr>
<td>Ghana</td>
<td>-0.093</td>
<td>+0.018</td>
<td>-0.135</td>
<td>-0.028</td>
<td></td>
</tr>
<tr>
<td>Nepal</td>
<td>-0.060</td>
<td>-0.109</td>
<td>-0.096</td>
<td>-0.132</td>
<td></td>
</tr>
<tr>
<td>Nicaragua</td>
<td>-0.094</td>
<td>-0.150</td>
<td>-0.124</td>
<td>-0.162</td>
<td></td>
</tr>
<tr>
<td>Pakistan</td>
<td>-0.051</td>
<td>0.000</td>
<td>-0.084</td>
<td>-0.017</td>
<td></td>
</tr>
<tr>
<td>Vietnam</td>
<td>-0.143</td>
<td>-0.009</td>
<td>-0.159</td>
<td>-0.016</td>
<td></td>
</tr>
<tr>
<td>Un-weighted Mean</td>
<td>-0.091</td>
<td>-0.064</td>
<td>-0.119</td>
<td>-0.076</td>
<td></td>
</tr>
</tbody>
</table>

Sources: Wagstaff, 2000b; Gwatkin, 2000b

The evidence presented by Gwatkin et al. (2000) also indicates that the distribution of utilization of health care services is extremely unequal across socio-economic groups. In spite of the fact that poorer segments of any society suffer disproportionally for death and morbidity, the wealthier segments utilize health services both in the private as well as the public sector disproportionally more than the poor. Even those services, such as interventions against

---

4 A third set of measures is presented by WHO (1999). The WHO estimates are not based on household data. They extrapolate the mortality rates at different incomes from information on the average mortality and on the distribution of income. The WHO estimates show stronger inequalities than those derived from household data. According to these estimates, those who live in extreme poverty (less than 1 US dollar adjusted for purchasing power parity) are five times more likely to die before reaching the age of 5 years than those above the poverty line, and two and half times more likely to die between the ages of 15 and 49 years.
communicable diseases, which address diseases mostly concentrated among the poor, are unevenly utilized more by the rich.

The following figure aggregates data for 43 developing countries on immunization of children\(^5\) for households belonging to different expenditure quintile groups. It shows that immunization services, which in principle would provide disproportionate health benefits to poorer households who are currently suffering from communicable diseases, are in fact utilized more by wealthier households. In Sub-Saharan Africa children belonging to the top quintile households are twice as likely to be vaccinated than those belonging to the poorest households.

**Figure 1: Inequalities in Immunization Rates between Wealthier and Poorer Households**

![Immunization Rates among Poor and Rich](image)

*Source: Soucat and Wagstaff, 2001*

The disproportionate use by the wealthier segments of society is due to better information, social norms more conductive to healthy behaviors (such as use of preventive services) and other demand side differences across socio-economic groups (see Annex 1). Hence, even if a new health program or service were to be made universally available (so that the poor would not suffer from higher financial barriers to entry, or need to travel longer distances), in general the rich would still end up utilizing the service more than the poor (see, for example, evidence on change in service utilization, after the introduction of a new Maternal and Child health (MCH) program in Ceara, Brazil, presented by Tendler and Freedheim, 1994). Thus, building a more even playing field on the supply side would not be enough to achieve equality of utilization.

Moreover, the evidence indicates that, given all the other determinants of health, even large disproportionate increases in the use of services by the poor would produce relatively smaller reductions in health disparities (for example, see evidence presented by Victora, 1999, on the health impact of a MCH program which carefully targeted the poor in the city of Pelotas, Brazil).

**THE DISTRIBUTION OF BENEFITS FROM PUBLIC EXPENDITURE ON HEALTH. EVIDENCE FROM DEVELOPING COUNTRIES.**

In this section we review recent benefit-incidence studies\(^6\) that focus on the benefits dispensed by public health systems to different socio-economic groups. These studies\(^7\) indicate that government expenditure on health care services and specifically on hospital care is pro-rich in most developing countries. A study by Gwatkin (2000b) summarizes the results of this literature as follows:

\[ X_j = \sum U_j \frac{S_i}{U_i} = \sum U_j \sum e_i S_i \]

**Mathematically, benefit incidence is estimated by the following formula:**

- \(X_j\) = health sector subsidy enjoyed by group j,
- \(U_j\) = utilization of service i by group j,
- \(U_i\) = utilization of service i by all groups combined,
- \(S_i\) = government net expenditure on service i, and
- \(e_i\) = group j’s share of utilization of service i

Thus, incidence studies do not consider therapeutic benefits, nor adjusts for different need across quintiles. They estimate the financial redistribution and do not consider the revenue side. They also do not convey any information about the reasons behind differential use of services across socio-economic groups.

---

\(^6\) For a review of the early literature, see Birdsall and Hecht, 1995.

\(^7\) Benefit-incidence studies compute the average subsidy for the different types of care (usually, primary, outpatient specialist care, inpatient care) from budgetary information and from data on service use. They estimate the subsidy going to each group from information on the distribution of utilization by income group of the different service-types.

---

7
Table 4: Percentage of Financial Subsidy from Government Health Services Accruing to Poorest and Richest 20% of the Population: Continent Averages

<table>
<thead>
<tr>
<th>Region</th>
<th>Primary Care</th>
<th>Hospital care</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Outpatient</td>
<td>Inpatient</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Africa</td>
<td>15 (7)</td>
<td>23 (7)</td>
<td>12 (2)</td>
<td>36 (2)</td>
<td>16 (2)</td>
<td>34 (2)</td>
<td>10 (5)</td>
<td>33 (5)</td>
</tr>
<tr>
<td>Asia</td>
<td>21 (2)</td>
<td>16 (2)</td>
<td>7 (1)</td>
<td>41 (1)</td>
<td>5 (1)</td>
<td>41 (1)</td>
<td>13 (1)</td>
<td>22 (1)</td>
</tr>
<tr>
<td>E. Europe</td>
<td>16 (2)</td>
<td>22 (2)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>12 (2)</td>
<td>29 (2)</td>
</tr>
<tr>
<td>Latin America</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

**Total Government Health Care Expenditure**

<table>
<thead>
<tr>
<th>Region</th>
<th>Poorest Quint.</th>
<th>Richest Quint.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td>12 (7)</td>
<td>30 (7)</td>
</tr>
<tr>
<td>Asia</td>
<td>19 (5)</td>
<td>21 (5)</td>
</tr>
<tr>
<td>E. Europe</td>
<td>13 (2)</td>
<td>27 (2)</td>
</tr>
<tr>
<td>Latin America</td>
<td>29 (8)</td>
<td>14 (8)</td>
</tr>
</tbody>
</table>

*Source: Gwatkin 2000b.*

Note: Each figure in parentheses indicates the number of countries included in the continent average that appears immediately to the parentheses’ left.

Note that according to the above table only in Latin America government expenditure on health care seems to be pro poor. In LA only Brazil shows a pro-rich distribution of public expenditure on government-run facilities. However, the estimates available do not take into account the public subsidy to social security, which covers only formal sector employees and that accounts for over half of total public health expenditure in several LA countries. In any case, inequalities are starker in the least developed countries of Sub-Saharan Africa. A benefit incidence study in seven countries in Africa shows that the richest 20% of the population receive well over twice as much financial benefit than the poorest 20% from all (inpatient as well as outpatient) government expenditure (Castro Leal, F., Dayton J., Demery, L. and K. Mehra. 1997). In all but two of the seven countries (rural Kenya and South Africa), the richest 20% also gain more than the poorest 20% of the population from primary care expenditure.

8 In general in LA public services fulfill mainly the residual role of providing basic services for those unable to pay, whilst the wealthiest segments of society utilize private facilities. Those employed as civil servants or formally in the private sector are frequently reimbursed by their own employment category’s insurance or sickness fund, frequently subsidized by government sources. Moreover, note that because benefit incidence studies do not take into account the revenue side, they tend to overestimate the redistributive impact of government health expenditure. In most developing countries, the main sources of financing it for the government are indirect taxation, payroll contributions and out-of-pocket payments, which are all regressive.
In Asia\textsuperscript{9}, Van de Walle (1998) in a study on Indonesia concludes that the redistributive impact of public health expenditure is negligible. Using benefit incidence measures, only in Malaysia and in the Philippines public health expenditure seem to be pro-poor (Castro-Leal, 2000). In a benefit incidence analysis on India, Mahal et al. (2001) show that the public subsidy for health is regressive, with the rich getting almost three times more than the poor, as the following figure indicates.

**Figure 2: Share of the Public Subsidy for Curative Care according to Income Group**

![Bar chart showing the share of the public subsidy for curative care by income quintiles.](chart.png)

*Source: Mahal et al., 2001*

Mahal’s study in India clearly indicates that different health services are characterized by an opposite redistributive impact. By measuring the cumulative benefits of different services for different socio-economic groups using a Lorenz curve, Mahal et al. (2001) show that the public subsidy to hospital care (inpatient and outpatient) is more pro-rich than the subsidy to outpatient care provided by primary health care facilities (PHC)\textsuperscript{10}, as the following figure shows:

\textsuperscript{9} Note that even if benefit incidence analyses in Asia show less unequal distributions than in Africa, the share of public expenditure over the total is on average much lower in the former than in the latter continent. A larger share funded by private out-of-pocket payments entails a larger burden on the poor, and overall health benefits more skewed in favor of the rich.

\textsuperscript{10} Within hospital care, inequalities across socio-economic groups are stronger for rural than for urban populations, suggesting that in India the rural poor cannot access hospital care. Also, stark differences exist across Indian states. Kerala is the only state with a pro-poor record for curative services. In three states, Maharashtra, Gujarat and Tamil Nadu, distribution of curative services is not statistically different from equality, whereas in all other states benefits’ distribution is strongly pro-rich, particularly in North and North East India (Uttar Pradesh, Bihar).
Note that the evidence concerning the distribution of health care leads to different conclusions in industrialized countries. These countries achieved a fairly even distribution of the financial burden and of utilization rates across socio-economic groups, although there are still pockets of under-served poor (Van Doorslaer and Wagstaff, 2000). Despite this, inequalities in health status across socio-economic groups persist and, in fact, during the 1990s they seem to have

Van Doorslaer et al. (2000) look at the distribution across income groups of utilization and need-adjusted utilization rates across 10 European countries, plus the USA. The paper highlights the following results:

Utilization of all services is pro-poor. Hospital use, which is pro-poor everywhere, influences this result, which however for some of the countries is not significant.

Utilization rates become less pro-poor and even moderately pro-rich once we take into account of increasingly precise indicators of need. Van Doorsaler et al. utilize age, sex, self-assessed health status (SAH) measures and a vector of chronic conditions to adjust for need.

Utilization is pro-rich for outpatient specialist care. The existence of a GP gate-keeper role does not seem to influence the rate of specialists visits, which remain pro-rich even in countries such as Denmark, Finland and the Netherlands, where patients do not have direct access to public specialist visits.

The presence or absence of universal public health insurance does not seem to significantly influence the distribution of utilization rates. The US and Switzerland do not emerge as outliers, and in outpatient care show similar pro-rich results as countries such as Sweden and East Germany, characterized by universal insurance and by a totally different payment system for providers.

For 6 countries, it has been possible to separately analyze GP visits. These turn out to be evenly distributed everywhere, but in Belgium and in Ireland, where they are pro-poor. It is not possible to determine whether the pro-poor distribution in Belgium and Ireland depends on the exemption of poorer individuals from outpatient care co-payments, or whether it is due to the impossibility of adjusting for SAH in Ireland and for chronic conditions in Belgium.
been widening in some countries (for the UK, see: Report of Health Inequalities, 1998). The increasing health inequalities observed in several industrialized countries are generally ascribed to the rising income and other socio-economic inequalities, and not to health system’s specific variables (see Annex 1).

Finally, note that the evidence from several developing countries indicates that government expenditure on health services is also very unequal in terms of geographical distribution within each country, with the wealthiest districts and areas being disproportionately favored. This is illustrated for Cambodia in the following Case Study 1.

**Case study 1: Distribution of health benefits in Cambodia**

A measure of general development is provided by the Human Development Index, which incorporates infant mortality, literacy, and GDP per capita into a single index (1 is the highest level possible, 0 the lowest). UNDP computed HDIs for each district in Bangladesh (UNDP, 1996). The relationship between the HDI and public spending per capita is illustrated in the figure below. The general trend suggests that districts with weaker development, as measured by the HDI, receive lower funding by the Government.

**Figure 4: Relationship between Human Development Index and Public spending per Capita in Bangladesh**

(1 BTaka = US$ 0.025-0.030. Bangladesh 1999-2000)

*Source: Bangladesh PER, 1999/2000*
WHAT IS THE MEANING OF AN EQUITABLE DISTRIBUTION OF
HEALTH AND HEALTH BENEFITS?

What do we mean by saying that a certain distribution of health, or health care services, or that a
specific repartition of the burden for financing health services are equitable or inequitable? The
literature that has analyzed this question focusing on health and health care draws upon an
immense literature in economics, sociology, political science and philosophy, which has tried
to address the issue of equity from a general standpoint.

Rather than just limit ourselves to presenting some of the measures that are currently utilized in
the health economics literature, in this section we chose to reconsider the notion of equity
starting from basic principles. This may be a roundabout approach to the central matter of the
paper, the equity impact of RAP reforms, and those who find this kind of conceptual discussions
uninteresting are encouraged to skip this section and go directly to the presentation of RAP
reforms.

However, we believe that attempting to present this discussion by broadening our perspective to
other sectors and by going back to basic principles may be useful for three reasons: a) to provide
the means to read with greater awareness the assumptions at the basis of the literature on equity
in the health sector, equity and reforms, etc.. b) To propose in a unitary, synthetic, and hopefully
original way several among the most significant contributions to the matter of equity and health
care produced in the last two decades. This is a literature which those working on applied health
system reforms are generally not familiar with. c) To frame the issue of equity in health care
within the government versus markets debate, and thus pave the way for understanding our
approach to analyzing the equity impact of RAP reforms, based on the notion of trade-offs. The
main points of the discussion are presented in the text, whilst further details are presented in
Annex 2.

The key question we try to address is the following: according to which criteria should resources
for health care be distributed among individuals?

12 Le Grand (1991, Chapter 5) provides a clear and insightful review of the main criteria advocated in economics,
equity as envy-freeness, as equality of income, of utility, or of marginal utility, and as absence of rank-reversal.

13 In modern times what characterizes a just or equitable society and a just distribution of resources, assets, etc. has
been a matter of contention since the early 17th century, when the Iusnaturalist philosophers put equity at the center
stage of the philosophical debate. They were so named because they believed in the existence of a "Ius Naturalis" a
Natural Law, which preceded and was superior to any earthly power. According to such Natural Law men were
created with some equal non-alienable rights. John Locke (1634-1703) was one of the first and more influential
Iusnaturalist philosophers. All modern positions based on the notion of human rights find their root in the Ius
Naturalis philosophy.

14 For a careful and clear review of the existing measures of health inequalities, both in utilization and outcomes,
see Wagstaff and Van Doorslaer, 2000.

15 Among all the articles and books on the issue of health and equity, I believe two contributions stand out for
clarity and thoughtfulness, and are thus recommended to those interested in further reading: Le Grand (1991) and
Culyer and Wagstaff (1993).
First, it is useful to clarify the distinction between equity and equality. As Le Grand (1991) correctly points out, the two concepts should be kept distinct. Equity, like justice, is a normative concept, whereas equality is an empirical, descriptive concept. In general, achieving greater equity does not necessarily imply moving towards more equality; the opposite may well be true\textsuperscript{16}.

Once the distinction between equity and inequality is clear, the answer to the above question stems directly from the analysis of how unequal the distribution of resources for health care is in our societies, compared to the “benchmark distribution” that we regard as just. A particular distribution is considered inequitable is there is discrepancy between the two and if its causes are amendable.

Hence, the first question to ask is the following: “What would be a just distribution for resources for health care? The answer to this question will depend on whether or not we believe that health is a fundamental human right (for a further discussion of this point, see Annex 2).

In general, the benchmark distribution for any given resource or asset is not perfectly egalitarian, and it can be different according to each of us. However, for everything that is considered as a human right, matters are different and more straightforward. By definition, individual human rights, unlike assets and resources, are inviolable, inalienable (one cannot buy or sell his/her right to vote, or his/her right to free speech) and must be equally enjoyed by all. In the literature more specific to health, Culyer and Wagstaff (1993), following Sen and a tradition deeply rooted in Europe, claim that health ought to be considered as a Positive Individual Right, and thus they conclude that equity demands that all equally enjoy it. The same position is echoed in recent official policy documents\textsuperscript{17}, and, for certain segments of the population, such as children, in international treaties\textsuperscript{18}. Thus, if health is indeed a human right, the answer to our above question is straightforward: resources for health care should be redistributed as much as possible to the benefit of those who currently cannot enjoy this right, or cannot enjoy it to the fullest\textsuperscript{19}.

\textsuperscript{16} Consider the example of a society where everybody is equal in terms of wealth, income, etc.. One may well argue that such society is inequitable because individual rewards do not reflect each person’s effort. Then increasing equity, as well as efficiency, in such a society would probably require increasing inequality. Note that the distribution of all resources, assets, etc is always unequal in our societies, although to a different extent. Judging whether this inequality is also equitable or inequitable requires a value judgment based on our perception of what would be the ‘fair share of the pie’ that each of us should receive. In general, knowing how the pie is divided is not sufficient. An example proposed by Le Grand (1992) describes the situation where two sisters, Mary and Anne, receive an equal share of a pie, even if only Mary helped grandmother to bake it. In such a situation, giving Mary a greater piece relative to Anne to reward her work would increase equity. However, suppose that Anne could not help because she was sick. Then, sharing the pie in half may be just. In other words, we cannot judge whether or not a situation is equitable just from information concerning end-states (equal or unequal division).

\textsuperscript{17} See, for example, WHO World Health Report 2000.

\textsuperscript{18} For instance, the Convention on the Rights of the Child, ratified by every country with the exception of the United States and Somalia, recognizes health among the fundamental rights of all children (Art. 24).

\textsuperscript{19} If we endorse the human rights perspective for health, then equity would demand that health outcomes be equalized (Culyer and Wagstaff, 1993), which, given the current distribution of burden of disease, would entail that the distribution of health inputs should be radically redistributed in favor of the poorer Nations, and in favor of the poor in each Nation.
Suppose we do not consider health as a human right. Then, to assess whether the distribution of health benefits is equitable or inequitable, one could refer to the criteria utilized for distributing other important resources, such as land. Here, all different positions may be ordered along a line, at the two extremes of which we find the following two. According to the first, the distribution of assets, resources, etc...in our societies is fundamentally just, and those few instances of unfairness are a matter of unavoidable chance. Therefore, in order to be just or equitable, any externally imposed change must preserve such original evenness. For example, a tax cut needs to be proportional and preserve the relative incomes and wealth, or at a minimum not to lead to rank reversal (see Feldstein, 1976), because that relative income and wealth truly reflect each individual’s relative ‘merit’. According to the opposite polar position, our societies and the existing allocation of resources in them are fundamentally unjust, largely the fruit of uneven initial conditions, privilege or “undeserved” luck^20.

Applying the two opposed views to the distribution of health, health resources and services at the global level and within each country, one may argue that the existing distribution, which is everywhere strikingly in favor of rich countries and of rich people within each country (see evidence presented before) is equitable, because access to health services should be part of any society’s reward system, or because the existing distribution simply reflects a high income-elasticity of demand for health services. Or, taking the opposite view, argue that the existing distribution is utterly unjust, because it reflects inequitable privileges and uneven initial conditions.

Rawl’s Theory of Justice (1971) provides an insightful example of the consequences of taking the second of the two above positions. Rawls claims that it is not possible to build a theory of justice based on people’s current position in society, because the latter is arbitrary, and because each individual would then choose different guiding principles to maximize her/his own benefit, given her/his position. According to Rawls, if instead people could choose under a “veil of ignorance”, that is to say without knowing which position they will end up occupying, they would unanimously opt for two principles, the first of which guarantees the maximum attainable level of liberties for all, and the second of which contains the Difference, or Maximin principle^21. Such principle states that inequalities may exist if and only if they are to the benefit of the poor. In case of health this would imply that we should stop re-equilibrating health care inputs, only when further intervening would create a negative impact on the health of the poor.

^20 For example, considering land-ownership distribution, which is extremely unequal in several countries, according to the first perspective one would argue that such distribution is just because it reflects individual preferences, effort and marginal productivity, whilst according to the second perspective the current distribution is utterly inequitable, the result of privileges consolidated over the course of several generations. According to this second view, an equity-enhancing change would demand a radical redistribution of land in favor of the land-less and the poor, whilst according to the first one should try not to upset the existing order, and just remove impediments to the ‘natural’ exchanges occurring through market interaction.

^21 They would choose such second principle out of self-interest and of fear of “later” being destined to occupy the place of the most disadvantaged in society. The economist philosopher Harsanyi showed that the Difference Principle could be chosen if and only if individuals were extremely risk-averse.
Note that in our societies assessing how equitable is a certain distribution of any asset or resource ultimately requires judging how equitable market mechanisms are in allocating resources. Markets and prices perform two main tasks in our economic systems. The first one is to co-ordinate the economic behavior of individuals. By providing signals of the relative scarcities of different commodities and assets, market prices are able to achieve allocative efficiency if certain conditions hold, as investigated for over one hundred years by economic theory. However, prices in reality perform another task: they ration scarce economic resources among individuals, according to their willingness and ability to pay. Given a certain price of a good or service, two individuals, of which the first is not willing to pay that price given her preferences and the second is not able to pay that price given her wealth and income, are treated in exactly the same way by the market. They are excluded from consumption of that particular good or service. The same logic applies to the interaction between countries. If -let us say- access to medicines is regulated by market mechanisms, a country who is not interested to introduce a specific drug and another who is not able to pay for it are treated in exactly the same way: they are both excluded from consumption of that particular drug.

In general, the literature on equity and health has taken strong egalitarian positions, refusing market allocation mechanisms as a benchmark to follow for health services, and advocating a strong re-equilibration of the distribution of resources for health care, both across countries and within each country in favor of the poor and those in more need. In the measures more widely utilized in this literature, such as the concentration index, any departure from the state of perfect equality in the distribution of resources for health is considered as inequitable. In assessing how much additional care over and above what the individual could him/herself afford or choose, this literature has presented several principles of equity, based either on equality of access, or equality of utilization, or (re)distribution of health care treatment according to need. In the following section we briefly present these principles.

By contrast, considering the issues of government failures and the limits to redistribution, some believe that the extent of market correction and of individual protection should be limited to a minimum “safety net”, which may include some essential services such as child care or emergency services, but that should exclude most other health treatments. These should instead be left to individual choices and market allocation criteria.

---

22 If certain conditions are satisfied (among these, price-taking behaviour by individuals and firms, perfect information, completeness of markets and absence of externalities and public goods), market prices determine equilibrium allocations that are Pareto optimal (first welfare theorem).

23 In other words, any of us is able to satisfy any need or preference he/she may develop for any commodity constrained by his/her Opportunity or Choice Set, which in our societies is largely determined by our “performance” according to market criteria.

24 For example, proponents of the Public Choice Theory (Brennan and Buchanan, 1977) underscore that it is not at all guaranteed that public intervention will achieve more equitable results than markets. According to this theory politicians and civil servants are in general self-interested individuals who utilize their public offices to consolidate their personal power and wealth, and their decisions hardly reflect their constituencies’ welfare.
**EQUITY IN HEALTH CARE. THE DIFFERENT PRINCIPLES IN THE LITERATURE**

**Equality of access and equality of utilization:** According to Le Grand (1982), potential equal access to treatment means that individuals should face the same money and time costs for using health services. Mooney (1983) defends such option as an equity benchmark because it is more respectful of individual preferences than the principle of equality of utilization or treatment: faced with the same prices, individuals may still choose different levels of treatment. However, Olson and Rodgers (1991) argue that equal access actually means that everybody can potentially consume the same amount of health services. Hence, prices must be the same (traveling time and costs included), given income, and the poor must be subsidized for their health consumption by the rich, until their opportunity cost for treatment is equalized. Pushing Olson and Rogers’s (1991) argument to its logic conclusion, Le Grand (1991) argues that equality of access ultimately requires that Choice (or Opportunity) Sets be the same for all individuals.

**Distribution of health care according to need.** Another commonly advocated equity principle for redistribution of health resources is need\(^{25}\). This view advocates equal treatment for equal need, and that those who are in greater need should come first.

But how shall we define need? In fact, Culyer and Wagstaff (1993) show that there are several possible definitions and measures of need:

- need as severity of illness;
- need as ability to benefit;
- need as the minimum amount of resources required to exhaust capacity to benefit.

Each of these definitions of need leads to different set of prescriptions for the distribution of health care services. Assuming for the purpose of illustration Quality Adjusted Life Years (QALys) as a measure of the quantum of health that each individual did or could enjoy\(^{26}\), we illustrate some of the possibilities of assuming the above definitions of need in table 5.

---

\(^{25}\) Usually, the literature refers to the principles of horizontal and vertical equity. The two principles were first articulated by the Greek philosopher Aristotle: the horizontal equity principle states that equals should be treated equally, whereas the vertical equity principle states that those who are unequal should be treated differently. Applied to health provision they are usually translated in the principle that access or use of health services should solely be based on need, and independent of socio-economic condition or of any other non-need related characteristic (such as gender). Applied to health financing they are usually translated in the principle that contribution to funding of health services should be in direct relationship with one's ability to pay.

\(^{26}\) The principle limitation of QALYs is that they attempt to measure quality of life on the basis of “expert” opinion, which can be arbitrary. The welfare (subjective) assessment of benefits, as well the social evaluation of health gains, may be totally different from any “objective” measurement of health gains.
Table 5: Equity and (re)distribution of health resources

<table>
<thead>
<tr>
<th>Item</th>
<th>Retrospective health</th>
<th>Prospective health</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Health from health care</strong></td>
<td>Previous QALYs gained from health care</td>
<td>Expected QALYs benefit from health care</td>
</tr>
<tr>
<td><strong>Health not from health care</strong></td>
<td>Previous QALYs enjoyed without health care</td>
<td>Expected QALY profile without health care</td>
</tr>
</tbody>
</table>

According to the ability to benefit exception of need, in this example we would advocate the criteria on the top-right cell to distribute health services. Equity would demand that expected QALY gains be equalized across individuals. The perspective is forward looking and would give priority to those most able to benefit from health care. There are two exceptions of the notion of “capacity to benefit”. One is to consider “marginal capacity to benefit (MCB)”. Equalizing MCB across individuals is the principle that maximizes incremental health benefits from health care. This is the most straightforward and widely used efficiency principle in health economics, but it is also sometimes advocated as an equity principle in the literature. In fact, it can lead to extremely inequitable choices, whenever those who are marginally more able to benefit (and should then come first according to this principle) are also the wealthier, the more healthy, or both.27 A second exception is Culyer and Wagstaff’s (1993) “minimum amount of resources required to exhaust capacity to benefit”. By these means, a consideration of severity of illness is indirectly phased in the definition of need as capacity to benefit, and the results obtained are in general more egalitarian than by just considering marginal capacity to benefit (see Culyer and Wagstaff, 1993).

However, equity criteria based on an exception of ‘need’ different from ‘capacity to benefit’ are possible as well, and they are indeed frequently preferred by individuals and households, when asked about priority setting criteria. In Sweden, for example, care for chronically, terminally ill and severely handicapped patients which ranked extremely low in terms of cost-effectiveness criteria were found to be popular and classified as ‘essential’ care in all household surveys that were conducted to inform the government on people’s preferences.28 Choosing the bottom-right cell, where need is interpreted as severity of illness, priority is given to those who would be more severely ill (would enjoy the “least number of QALYs”), without care, regardless of their ability to gain extra health as a result of this care or their ‘objective’ capacity to benefit.

Another principle would be to use the same criteria accepted for education, where everybody is given a set amount of free services (equality of inputs), and every extra is given according to merit, or it is fully paid for. In the case of health, one should then equalize care received across individuals. Alternatively,

27 This criterion ominously recalls principles originally proposed by Herbert Spencer. Spencer (1820–1903), a British sociologist, first applied Darwin’s theory of evolution to the study of human societies. According to his views, individuals who contribute more to a society and those who are more fit should be preferred over the others.

one could try to equalize health benefits, or QALYs gained from subsidized care\textsuperscript{29} over the life cycle, and choose according to the criterion in the top-left cell, table 5, column 2. Finally, one might want to bring in considerations linked to the age or the past health condition of potential beneficiaries and give priority to the young or to the disabled (those who have enjoyed the least QALYs without health care), according to the criterion in the lower-left cell of table 5, column 2.

Note that all of these equity-motivated criteria, with the exception of the marginal capacity to benefit-based redistribution principle, also conflict with any possible effectiveness principle, which would tend to put resources first where the link between inputs and outcomes can be stronger. In other words, there is in general a trade-off between distributing resources for health care to maximize gains in health outcomes, and to achieve greater equity (equity-efficiency trade-off).

All of the above discussion should suggest how complicated any equity assessment is. By adopting any of the above different principles we may be lead to opposite conclusions concerning the impact on equity of the same change. Case study 2 below illustrates a concrete example where, by adopting a different principle and consequently a different measure of the degree of equity or inequity in the distribution of health care resources, we would come to a different assessment of the impact in terms of equity of the same health policy intervention.

**Case study 2: The equity impact of social health insurance in Ecuador**

Waters (2000) estimates the equity impact of a new social insurance program which covers health services for formal sector employees in Ecuador, and of a plan for its extension to self-employed or to the workers’ dependants. He points out that many of the measures currently used to measure “inequality” (see for example, Wagstaff and van Doorslaer, 2000) negatively assess any departure from a state of perfect equality in health status or in need-adjusted utilization of health services, without consideration for the “levels” of care achieved by the different groups. Thus, the measures of inequity generally adopted in the literature imply that any pro-rich departure from a state of perfect equality is inequitable, even if the poor are positively affected. By using two of these “egalitarian” measures of equity\textsuperscript{30}, the concentration index and the Atkinson distributional measure\textsuperscript{31}, Waters (2000) shows that distribution of utilization of health services in Ecuador was already pro-rich, and it was made even more inequitable by the new social insurance program. The new program would still be an adverse impact on equity even if were extended to

\textsuperscript{29} The issue here is not to decide priorities starting from a situation where no services are available; rather, to evaluate how much additional care should the state allow an individual to receive over and above what he/she could afford to pay for.

\textsuperscript{30} Let us briefly explain a general formulation of the existing measures of inequity, based on the measurement of inter-group or inter-individual differences in access to health. If \( H_i \) is a measure of access or health status of socio-economic group \( i \), where \( i=1,\ldots,I \) ranks individuals or groups in increasing order according to their socio-economic status (for instance, their income or expenditure), the degree of inequity is measured by a direct function \( f \) of the rich-poor differences: Inequity = \( f \left( \sum_i \sum_{j>i} (H_j - H_i) \right) \) with \( j>i \). In the need-adjusted measures based on utilization of services, the above formula is replaced by: Inequity = \( f \left( \sum_i \sum_{j>i} [(H_j | N_j) - (H_i | N_i)] \right) \), where \( H_i \) is a measure of utilization of services and \( N_i \) is a measure of health need (for instance, self-reported health status) for socio-economic group \( i \). Paradoxically, using such measures of inequity, a scenario where no person in need could receive any health care would be considered more equitable than a scenario where the poor receive good care, but inferior to that received by the rich.

\textsuperscript{31} The Atkinson distributional measure evaluates the total welfare loss from an unequal distribution of utilization of health services.
dependants or self-employed. The benefits of the social health insurance program would still be regressive, although less so in case the program were extended to the self-employed or to dependants.

However, Waters also considers an alternative view of equity, according to which an intervention that improves coverage for the poorer members of society could be considered as equity-enhancing, even if increased the level of overall variability in health benefits across the population. As a third measure of “inequity” in line with this second view of equity, Waters proposes a weighted welfare function, where the utilization of services by the poorer quintile is weighted more than for the second-poorest quintile, etc. Using this other measure of the degree of inequity in the distribution of health benefits, Waters shows that equity would be unequivocally enhanced by the social insurance program, particularly if the benefits were to be extended to the self-employed or dependants. In conclusion, he shows how different notions of equity in the distribution of health benefits can be translated into different measures, and how these would lead to a different evaluation of the impact on equity of the same interventions.

The conceptual literature on health and equity was developed mostly in Western Europe. If we compare the concrete evidence on the distribution of health and health benefits in developing countries presented in the first section with the above conceptual discussion, the contrast between the two is palpable. For example, much of the above conceptual literature on health and equity is centered around accessibility problems, in a sense giving for granted that health services, once made accessible for all, would also be of acceptable quality standards. By contrast, by looking at the evidence from developing countries we see that the quality dimension of services is of great importance (see following section). In several countries or regions within countries public health services are in theory accessible to (almost) all, but they are of such abysmal quality particularly in the peripheral and poor areas that in reality very few utilize them. In these settings, a focus on equality of treatment, or equal treatment for equal need is more appropriate than just focusing on accessibility.

Second, in analyzing how the current distribution of health benefits can be concretely improved in developing countries, one has to take into full account the political economic and other constraints under which developing countries’ governments must operate. For example, all the above criteria advocate that priorities ought to be set among individuals and not services, but at best what developing countries’ governments can do is only to prioritize across services rather than individuals, because of the impossibility of assessing people’s need or socio-economic status, or because of the political or technical impossibility to exclude certain individuals from publicly subsidized services\(^\text{32}\). Considering these political-economic constraints, some claim that the best way to reduce health inequalities would in fact be to expand the scope of universal programs. However, the “mapping” from services to individuals is far from perfect: without explicit targeting, most health benefits end up being captured by the better off, as indicated in section 2.

\(^{32}\) For example, it may be impossible to enforce the provision of a basic package of services exclusively for the poor, because doctors and other health workers, who ultimately make the rationing decisions, are not willing or able to comply with it.
WHAT IS THE MEANING OF RAP REFORMS?

If one tried to specify the two terms of an equation to capture the equity-RAP relationship, a measure of equity, on the one hand, and an indicator of the change in resource allocation, payment systems and purchasing agreements on the other, one would encounter into all sorts of conceptual as well as measurement difficulties. We presented alternative definitions of equity in the previous section. We now turn to the right-hand side of the equation, and define what are the core elements that characterize RAP reforms. Given the complexity of the matter, we will follow a step-by-step approach and we will by necessity include only a selective review of the empirical literature.\(^3\)

RAP mechanisms define the criteria according to which funds, collected through different revenue sources and pooled together, flow within the health system, eventually reaching service providers.\(^4\) The place of resource allocation and of purchasing arrangements in relation to the other dimensions of health financing and service delivery can be visualized in Figure 5, taken from Preker et al. (2001).

Figure 5: Financial Flows Within a Health System

![Flow of Funds Through the System](image)

Source: Preker et al., 2001

---

\(^3\) We hope that, although being incomplete, this introduction to the meaning of RAP reforms and their impact will resonate with others’ experience, and stimulate further interest and discussion.

\(^4\) Note that health financing includes both the alternative ways to fund health services, by collecting and pooling resources for health, as well as the criteria for distributing resources across purchasers and providers. RAP reforms are meant to change the latter, and not the former.
The so-called RAP reforms have both: a) strengthened RAP mechanisms and b) introduced a multidimensional vector of changes to existing RAP arrangements. We will focus on six key components, five influencing the supply side, and one on the demand side. The six RAP reforms components are:

- redesign of resource allocation criteria, moving from input-based, retrospective towards capitation-based, or other forms of prospective financing;
- redesign governments’ priorities across services or levels of care (in several developing countries this component of RAP reforms has been identified with the introduction of the “Basic Benefit Package, or BBP);
- the interposition of a purchasing agency between patients and providers. Such agency is generally entrusted with the role of commissioning/contracting services from (semi)independent providers;
- change of payment criteria for providers, from input-based towards prospective or activity-based payment systems;
- provision of explicit financial and other incentives/enablers to providers, linked with their ability to reach the poor/cure diseases that primarily affect the poor;
- financial incentives for patients/clients, such as vouchers, to stimulate consumption of specific health care services, such as prenatal care and institutional delivery.

RAP reforms implemented over the last 15 years in several industrialized and developing countries included a different combination of the above six elements. Table 6 synthetically illustrates the above RAP reform components, their principal objective, purpose and activities.
<table>
<thead>
<tr>
<th>RAP reform component</th>
<th>Priority setting</th>
<th>Resource allocation criteria</th>
<th>Purchasing/contracting</th>
<th>Providers’ payment systems</th>
<th>Explicit incentives for providers to reach specific groups/cure certain diseases</th>
<th>Explicit incentives for patients to use specific services</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAP arrangement addressed</td>
<td>What services to subsidize and what to exclude?</td>
<td>What are the rules (formula) for transferring public funds?</td>
<td>To make or to buy decision</td>
<td>How to pay? How much to purchase and at what prices?</td>
<td>Whom to purchase for?</td>
<td>Whom to purchase for?</td>
</tr>
<tr>
<td>Purpose</td>
<td>Strategic definition of priorities for public financing and provision.</td>
<td>New criteria for allocating resources across regions, districts and purchasers.</td>
<td>Strategic purchasing; more transparency and results-orientation in public sector; engagement of the private sector in service provision.</td>
<td>Redefinition of the incentive regime for providers.</td>
<td>Reduction of exclusion of the poor from public services/sharp er targeting</td>
<td>Increase utilization of essential services by the poor</td>
</tr>
<tr>
<td>Activities</td>
<td>Redefinition of priorities across services according to cost-effectiveness or other criteria.</td>
<td>From retrospective to prospective resource allocation.</td>
<td>Definition of contracts and service agreements, increased degree of market exposure for providers</td>
<td>Output/performance based payment systems for providers; performance-based incentives for health workers. Capitation based funding of primary care</td>
<td>Equity-motivated risk-adjusters to correct capitation funding. Explicit financial incentives for providers to cover the poor, or to provide care against diseases that affect the poor (ex. TB).</td>
<td>Specific enablers or incentives (such as vouchers) that contribute to treatment and traveling expenses</td>
</tr>
</tbody>
</table>

**Objective**
- Max impact on disease burden, cost effectiveness, correction of market failures
- Allocative and productive efficiency, Consumer-responsiveness
- Equity, Financial protection
As Table 6 indicates, a correct way to understand RAP reforms is to view them as a set of instruments, or a tool-kit, which can be varied to achieve improvements in RAP arrangements to enhance efficiency, quality, or improve accessibility of services for the poor.

Also note from Figure 5 that RAP reforms do not intervene on the sources of financing, or in the organization of service provision, although RAP reform components have frequently been introduced in the context of wider reform processes, which affected also these other dimensions of the health care sector (see case study below, on RAP reforms in Colombia). Specifically, the link of RAP mechanisms with the sources of revenue and pooling arrangements is crucial: any RAP arrangement presupposes that a health financing mechanism is in place to harness private resources for health care in a pooled fund. In other words, RAP criteria are relevant where governments/social insurance funding is a non-negligible part of total health funding, and where governments can decide how to allocate geographically. However, they are much less relevant to countries with prevalence of out-of-pocket payments, or demand side funding (e.g. patients choosing where to go and facilities simply being reimbursed retrospectively), unless some geographical or population group budget constraint can be imposed.

Case Study 3: Evidence from Colombia

For a few Latin American and former socialist countries we dispose of information concerning the evolution over the last few years of utilization of health services and formal insurance coverage by socio-economic status. In this section we review the evidence from Colombia, a case that is particularly interesting, because it was able to combine in a rational framework several key reform elements. Health reforms in Colombia included a resource mobilization, as well as several RAP components. Once again, we must say that such evidence concerns the impact of a whole, large-scale reform programme, and that it is really impossible to ascribe any of the changes observed to any particular component of RAP reforms.

The 1993 Colombian health reforms strengthened the health revenue collection and pooling systems, and they changed the criteria for allocating resources within the health system. The core RAP reform components were the following:

35 This is a key point: what RAP reforms have been able to achieve may be constrained by the fragmentation of the health funding system. For example, in former socialist countries there have been five parallel coexisting financing systems in place, each of them responsible for funding services for different population subgroups and/or for different services: (a) those in formal employment are often covered through contributions to a social insurance fund; (b) the poor and destitute are often covered by direct transfers from the Ministry of Health. The MoH also covers public health services (c) regions, districts and municipalities finance specific services, such as emergency services, through local taxes and through transfers from the Ministry of Finance (d) the rich sometimes have access to purchasing arrangements provided through private health insurance: and (d) rural populations and the informal sector rely heavily on direct out-of-pocket payments outside formal purchasing arrangements. A similar fragmentation characterizes also most countries in Latin America, where there are usually several sickness funds with significant differences in the benefits provided. In these contexts RAP mechanisms risk to be largely irrelevant, and disparities across socio-economic groups in the level of health benefits are likely to persist or increase over time.

36 This Case Study draws on a paper by Londono, Jaramillo and Uribe (1999).

37 On the resource mobilization front, the government raised the payroll tax rate on formal sector employees from 8 to 12 percent (the same rate holds for self-employed workers who report a salary above the national minimum), and hypothesized for health a fixed share of local governments' revenues (25% of the total local governments’ revenue, according to Law 60 of 1993). Such revenue mainly consists of transfers from the central government, plus minor local taxes.
The creation of a national contributory insurance scheme for formal sector employees, managed by Empresas Promotoras de Salud (EPS), and of a national subsidized insurance scheme, managed by Administradoras del Régimen Subsidiado (ARS). The EPS and the ARS are entrusted with the purchasing role: they receive a capitation funding for each beneficiary (la Unidad de Pago por Capitación – UPC), which is larger for children under 5, for the old and for women in fertile age, and they guarantee coverage of, or directly provide a prefixed package of health services to their beneficiaries. In prospect, the benefit package should become equal for all Colombians (Plan Obligatorio de Salud, POS), but so far it is more generous for those in the contributory regime than for those in the subsidized regime (approximately twice as expensive per capita, US$127 as opposed to $68 per capita in 1998).

The creation of a national compensatory fund (Fondo de Solidaridad y Garantía, FOSYGA), entrusted with the role of supervising the financial flows in the system, of cross-subsidizing EPS in deficit with resources from those in surplus (Subcuenta de Compensación), and of providing parts of the subsidies for ARS (the other part -69% in 1999- is provided by municipal governments) 38.

The transformation of hospitals and other health providers into independent Prestadoras de Servicios de Salud (IPS), legally regulated as ESE (Empresas Sociales del Estado) and mainly funded by EPSs and ARSs according to the services provided to the insured population. Thus, the reforms foresaw that the direct subsidies to providers, previously assigned mainly by local governments, should have been progressively phased out and substituted by activity-based payments, according to the services rendered to the insured population. However, recently the central government stopped the process of phasing out of the subsidies.

The evidence available from Colombia after 1993 shows weaknesses and gaps in the reform process. Nonetheless, overall it shows an overall positive impact of the health reforms on equity. In particular, some key macro-indicators of equity seem to have improved, in spite of the several difficulties encountered in the reform process, and in spite of socio-economic hardships the country continued to experience while the health reforms where in progress39.

Looking first at the changes in social health insurance coverage, the evidence shows that the Colombian health system before the reforms was extremely fragmented and inequitable, with marked differences in access to health services and their quality40. As a result of the weaknesses of the pre-payment schemes, the poorest decile of the population was spending on health a share of their income up to ten times higher than the richest decile, and approximately half of those who reported illness did not seek care because they were unable to pay41.

After the reforms the number of people insured through both the contributory as well as the subsidized regime has sharply increased, raising the total number of people insured from 23% to roughly 60% of the total population. The improvement was particularly significant for households belonging to the poorest expenditure quintile in the country, who increased their health insurance coverage in the subsidized regime more than tenfold, and those belonging to the second quintile approximately six times. In year 2000, of the

38 The central government should also match FOSYGA’s subsidies, a mechanism known as paripassu’. However, due to the difficult fiscal situation, such contribution was frozen.

39 The growth rate in Colombia has been negative for most part of the last decade, unemployment raised from 7% in 1994 to 22% in 1999, and the civil war created hundred of thousands of displaced people.

40 Social security covered health risks for approximately half of the workers in the formal sector; among these, only 20% enjoyed health coverage extended to their families. Overall, 72.2% of the adult working population in Colombia (and 90.8% of the poorer decile) was excluded from any social security benefits. Only 23% of the total population (approximately 7 million Colombians) enjoyed any form of health insurance coverage.

41 Molina C.G. et al. (1993).
total poor population in the country, 35.3% was covered by the subsidized regime, 10.7% by the contributory regime, and 53.9% as yet did not have health coverage. Other major achievements of the reforms were the following: A) Between 1993 and 1997, health subsidies for the poorest quintiles of the population increased by 200%; and those for the second poorest quintiles by more than 100% (Sanchez and Nunez, 1999). Meanwhile, health subsidies for the two quintiles of the most affluent sector of the population decreased 70% and 14% respectively. B) Disparities in the allocation of resources for health across regions and departments have been reduced with the process of decentralization and with the introduction of capitation-based resource allocation criteria. C) Most importantly, utilization of services has increased across all income groups, as the following table shows:


<table>
<thead>
<tr>
<th></th>
<th>Ambulatory Consultations</th>
<th>Hospitalizations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1993</td>
<td>1997</td>
</tr>
<tr>
<td>Quintile 1</td>
<td>340,856</td>
<td>605,840</td>
</tr>
<tr>
<td>Quintile 2</td>
<td>436,223</td>
<td>827,147</td>
</tr>
<tr>
<td>Quintile 3</td>
<td>549,161</td>
<td>1,232,268</td>
</tr>
<tr>
<td>Quintile 4</td>
<td>590,331</td>
<td>1,384,333</td>
</tr>
<tr>
<td>Quintile 5</td>
<td>598,863</td>
<td>1,362,215</td>
</tr>
<tr>
<td>TOTAL</td>
<td>2,515,384</td>
<td>5,411,803</td>
</tr>
</tbody>
</table>

Source: Londono, Jaramillo and Uribe, 1999

**Motivation of RAP Reforms**

RAP reforms are an attempt to respond to the issues of efficiency, quality, and equity plaguing health systems in several industrial, transition, and developing countries. Such issues, due to a combination of government and market failures, are revealed in poor spending choices, distorted allocation or resources, provider unresponsiveness to clients, and in general poor quality of services, etc., and ultimately, the inability of the health care system to contribute to achieve better health outcomes.

Over the last decade, a number of studies indicated that "bypassing" of low-quality services offered by public facilities is a widespread phenomenon (Akin and Hutchinson, 1999. For Kenya, see Mwabo, 1993; for the Dominican Republic, Lewis et al., 1992, and for Pakistan, PIEDR, 1994). The high price-elasticity of demand for public health services, particularly observed for child services and among the poor (for Uganda, see Giusti, 2000 and for Indonesia, Gertler and Molyneaux, 1990) may conceal a low evaluation of such services.

---

42 The Corporation for Research and Development (CRD), and the Medellin Economic Faculty (CIDE) found that, between 1987 and 1995, regions that at the beginning of the period had relatively less public resources, witnessed a greater increase in funds available, and that as a result inequalities in per capita available by department have decreased. Bossert (2000) confirms the same results.

25 Adapted from Sanchez F. y Nunez J. (1999). The numbers presented do not correspond to the total of the services offered in the country. They correspond to the estimations of the CASEN (1993) and of Quality of Life (1997) Surveys.

43 In 1997 a natural experiment was recorded in Uganda. Non-profit facilities raised their user-fees in August (up to roughly US $ 1.5 per child outpatient visit, and up to US $ 7.5 per adult outpatient visit) and then lowered them again in November, after the government committed to subsidize them. Giusti (2000) recorded extremely wide variation in service use, especially for childcare.
The evidence indicates a high degree of absenteeism in government facilities (see for example, results from study in India by Chaudhury and others, presented in WDR 2004, Chapter 2, Table 1.2b), dual job holding, widespread levels of corruption, etc., and it concludes that the level of health expenditure and health inputs officially recorded, such as the number of doctors or number of health facilities available in any particular area, create a negligible impact on health outcomes.

Note that these findings radically question the current effectiveness of public services, but do not question the effectiveness of health services per se. In fact, the evidence indicates that real health inputs and service quality produce a strong impact on utilization and on outcomes (Alderman and Lavy, 1996). For instance, Lavy et al. (1996) show that in Ghana drug availability and weekly hours of availability of physicians produce a strong impact on child survival, on child height and weight for height.

These studies also question the effectiveness of the traditional ways of channelling foreign assistance in the health sector to developing countries, mainly focused on the expansion of the infrastructure base of the government health sector. While these continuous investments have made a tangible difference to physical infrastructure and delivery capacity, the outcomes in terms of actual improvement in services – particularly to the poor and disadvantaged groups – have not been as obvious 44.

To draw a coherent ‘stylized fact’ from all of the above findings, we may say that increasingly it is investments in service improvements and not availability of facilities that seems to matter most. Furthermore, by looking at the variations across countries and at the few success stories available, it is becoming recognizable that the results obtained by a few countries in terms of health outcomes and in terms of quality and accessibility of essential health services are achievable also at relatively low levels of economic development. As indicated before, the international evidence indicates that the relationship between average national income (measured by per capita GDP) to priority health outcomes is by no means universal or automatic, and that low income by itself does not explain the poor health achievement of citizens of some developing countries.

A renovated public sector could play an important role especially in improving the health of the poor. In a study based on a sample of 35 countries (where nationally representative household surveys were conducted), Bidani and Ravallion (1997) show that public health spending matters for the poor (measured as those below US$2 or US$1 per day at 1985 purchasing power parity), although its impact is insignificant on the richer segments of the population and on the population as a whole. More recently, Gupta et al. (2001) reach similar results. These results are

44 A similar point is stated in the 2001 HNP Chapter for the PRSP Sourcebook: “One point needs emphasizing, namely that funds linked to PRSPs – whether debt relief or IDA credits - will have a far greater impact on poor countries’ health levels if they are accompanied by a thorough review of existing policies, and by a willingness to link new spending with reforms that make health systems work better, especially for the people they tend to serve less well – the poor (2001 HNP Chapter for PRSP Sourcebook, p.5). The Commission on Macroeconomics and Health (WP 5) reached similar conclusions in its work on constraints to scaling up health services (Mills et al. 2001).
not surprising, if we consider that the poor are less able to substitute private care for public care, when the latter is lacking or is of extreme poor quality\textsuperscript{45}.

Mahal et al. (2001) confirm similar results for India (see Figure 6 and 7) by showing that, even if public health expenditure is strongly pro rich, the poor still rely on the public sector more than any other socio-economic group.

**Figure 6: Public and Private Sector Hospitalization Rates by Income Quintile**

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure6.png}
\caption{Public and Private Sector Hospitalization Rates by Income Quintile}
\end{figure}

*Source: Mahal et al., 2001*

\textsuperscript{45} By looking at immunization, attended deliveries and other services, Gwatkin (2000b) shows that the negative outliers in terms of distribution of utilization are countries such as Turkey and Morocco, that mainly rely on private financing and provision.
Against the background described above, RAP reforms can be easily justified. Their focus is on the “value for money” that patients obtain from the health system, and to find ways to improve it. If the reason for poor quality and ineffectiveness in service delivery particularly in the public sector is to be found in the lack of incentives, or to the “wrong” incentives existing for providers, then a precondition to improve “value for money” is to profoundly reform the criteria according to which resource allocation decisions are taken, priorities are set, and financial rewards/penalties are given. RAP reforms propose to improve health outcomes by incorporating a more transparent and results-based approach as a key feature in their design.

Hence, the test to assess RAP reforms is whether such reforms are actually achieving the service improvements they promised or not, and whether the poor are at least partially benefiting from such improvements.46 Also, note that there is more to improving access and quality of services than just allocating resources more transparently and sensibly, and that therefore RAP reforms are but a piece in a large puzzle. Even if a correct allocation of resources, better incentives were set in place, we cannot assume that this will mechanically lead to the desired objectives. One other piece, probably even more important than RAP mechanisms as far as equity is concerned, is the revenue collection47 and pooling dimension. Here, we focus on measures aimed at

46 Note that RAP reforms have not been justified because they intended to improve services specifically for the poor. However, the above discussion suggests that any health reform that succeeds in improving quality and effectiveness of health services in general, and particularly public services, without reducing accessibility, it is also likely to have ipso facto a positive impact on the poor, because it is the poor who mostly have to rely on the government for financing their health expenditure.

47 The principal revenue collection systems are: (a) Out-of-pocket payments; (b) Voluntary insurance contributions (usually risk-related) with competing insurers; (c) Tax-based public funding; (d) Mandated employment-based insurance. The prevalence in the funding of any health system of one of the above funding systems determines the overall efficiency/equity impact in the distribution of the health financial burden. International evidence from
modifying the criteria according to which health resources are allocated and purchased, and we are not analyzing the impact of different revenue collection mechanisms\textsuperscript{48}. But the link between RAP and sources of revenue and pooling mechanisms is crucial, as we already mentioned. Each RAP mechanism is embedded within a specific revenue collection or resource mobilization system, which determines the limits of what the RAP mechanism can achieve also in terms of equity.

**THE RELATIONSHIP BETWEEN RAP REFORMS AND PRIVATIZATION**

Our brief characterization of RAP reforms should have already cleared some of the major motives of confusion commonly found in discussions concerning RAP reforms. In conclusion of this section, it is worth clarifying also that RAP reforms are extremely different from privatization of the health system. This is despite a number of facts. It is true that through purchasing and contracting, RAP reforms try to replicate within the public sector some of the mechanisms of the market for achieving greater efficiency. It is also true that the essential motivation of RAP reforms is the recognition that “incentives apply to governments and not just to markets” (Stiglitz, 1999, pp.3-4). Thirdly, it is true that with the purchasing reforms, governments have for the first time recognized the role of the private sector in health service delivery. Finally and unfortunately, RAP reforms in several countries have been associated with drastic cuts in public expenditure and in public services, rising user-fees, and privatization. However, it is important to maintain the distinction between privatization and RAP reforms. In fact, one of the more difficult challenges which RAP reforms confronted in several countries was to be able to grasp some of the advantages of markets in terms of efficiency and quality of services, and yet to preserve, or to enhance other goals, such as equity, that should characterize public health systems.

**RAP Reform Components: A Review of the Evidence of Their Impact on Equity**

In this section we briefly describe each of the six RAP reforms components introduced before, we selectively go over the experience from industrialized countries, and review the evidence on their impact in developing countries, with a particular focus on equity. In doing this we follow the flow of money within the system, and we go ‘downstream’, starting from the resource allocation and the priority setting components, and then turning to the analysis of the payment system and purchasing components, and finally the demand side incentive mechanisms (see figure 5).

industrialized countries shows that the only progressive funding mechanism is direct taxation. Employment-based social insurance contributions are generally proportional or mildly regressive, whilst private insurance contributions tend to be regressive (Van Doorslaer, 1997). Within a funding system exclusively based on voluntary insurance premiums, cross-subsidies across risk/income groups are extremely hard to maintain due to the well-known risk-selection and adverse selection problems. The only way the government can pursue equity-motivated corrections to market outcomes is through direct subsidies to the poorer and/or higher risk individuals in the first case, or to health plans that accept to offer coverage to the poorer, and higher risks in the second case. All pooling schemes based on individual voluntary participation will always be highly constrained in the amount of redistribution they can achieve precisely because they demand universal consensus of participants. In the language of welfare economics, they will promote only cross-subsidies that lead to Pareto-improvements.

\textsuperscript{48} For an insightful analysis, see Wagstaff (2001).
Reforms of the Resource Allocation Criteria

‘Resource allocation’ (RA) criteria define the flow of financial resources from the centre to decentralized levels of government (such as regions and districts) and to health purchasers. Sometimes the term ‘resource allocation’ is also used to indicate explicit priority setting policies that ring fence resources to stimulate the provision of specific services, but we will deal with these latter ones separately in the next section. RA criteria are closely related to payment systems, a term usually utilized to indicate reimbursement criteria for service providers.

Over the last two decades, several of the industrialized countries implemented RA reforms in the health sector, as well as other sectors such as education. They shifted from mainly retrospective criteria (based on historical allocations), towards population-based, prospective RA formulas for financing local levels of government and health purchasers. According to Rice and Smith (2000), the main motivation driving those changes has been the achievement of tighter cost-control, but the new RA criteria have in fact also introduced a greater transparency and equity in the distribution of public funds. The capitation formula has been in several cases complemented to take into account the relative health needs of different segments of the population. Some services deemed essential are also often ‘top-sliced’ (assigned to national vertical programs before allocation of the rest to lower levels), and not included in any allocation formula. Both individual (such as age and gender), as well as social and geographic factors have been extensively utilized in the computation of the RA formula (the so called risk-adjusters). For instance, the RA formula implemented in the ‘90s in Stockholm, Sweden, to distribute funds across different health authorities in different neighborhoods included as risk-adjusters age and gender of the resident population, housing tenure, marital and employment status, as well as previous utilization of health services.

The experience from industrialized countries, such as the UK or Sweden, suggests that:

a. Population and need based funding creates a great opportunity to equilibrate distribution of resources across regions and districts.

b. Devising a gradual, smooth transition process to the new regime, so that the ‘losers’ in the redistribution process have the opportunity to adjust and the ‘winners’ to develop new planning capacity is a key but problematic issue. In this respect, in the UK, the Resource Allocation Working Party (RAWP) approach is widely perceived as a success. In the ‘70s a new capitation-based funding was adopted, which lead to a progressive rebalancing of the distribution of resources across the country, by assigning any real growth in the budget to under-resourced areas. Previously, the Southeast region and London had received a disproportionate share of the total. The transition period was set up so that if a region was over-resourced its budget could be cut by 2.5% maximum per year, whilst if a region was below target, it could receive an additional maximum funding equal to 5% per year.

49 For instance the share going to services belonging to the Basic Benefit Package (see following section), or that going to tertiary care and teaching hospitals.
Relatively few of the developing countries underwent comparable reforms of their resource allocation criteria, and the majority of them still allocate public resources across regions, districts and health plans on a historical-incremental basis, with some adjustments achieved through political negotiation. The lack of population-based data concerning individual and social circumstances, as well as utilization of services, contributes to severely limit the variables that can be considered in the capitation formula.

Latin America is the region that has gone further in experimenting new population-based RA mechanisms. Countries that introduced significant changes in the resources allocation criteria include Brazil, Bolivia, Chile, Colombia, and Mexico. In Colombia, the government introduced a new RA mechanism based on a capitation formula for districts and municipalities, and adjusted it according to the share of the poor in each municipality and the fiscal effort at the local level. Brazil adopted a population and need-based formula to determine transfers to the states earmarked to finance primary care infrastructure, equipment and other capital expenditure (see following case study).

**Case Study 4: Resource allocation criteria for capital expenditure in Brazil**

In Brazil, the formula used to decide the extent of capital investment resources given to each state was specified as a weighted average of the state’s share of the total population, and an indicator of the relative health need of the state vis-à-vis the country as a whole. The formula was specified as follows (Musgrove, 1996, p. 3):

\[ F_e/F = a(P_e/P) + (1-a)(X_e/X) \]

where \( e \) refers to a particular government level, which, in the Brazilian example, was the state level (Estado); \( F \) is the investment, or other funds to be pre-allocated; \( P \) is population; \( X \) is a synthesis indicator of need for additional investment, beyond that associated with population alone \((\sum X_e = X; X_e \geq 0)\); \( a \), where \( 0 \leq a \leq 1 \), is the weight given to the population, as opposed to that assigned to the need indicator \( X_e \).

Initially \( X \) was computed simply as the inverse of total recurrent public expenditure in each state (labeled \( G_e \) for ‘gasto’), based on the assumption that where little was spent there was a greater need of further investment, and on the lack of data that could be used as alternative measures of need. However, using total recurrent public expenditure instead of per-capita expenditure, led to grossly overestimate ‘empty’ states’ need. For example Amapa, which accounts for 0.21 percent of the total Brazilian population, received about eight times more than it would have received based on population alone. Later on, relative recurrent expenditure per capita replaced total expenditure in the computation of \( X_e \).

The primary aim of the above prospective formula has been to achieve greater cost-control over local spending, to bring greater transparency, and to reduce inequities in capital investment across sub-national units. States such as Parana, which used to receive a disproportionate share of total capital expenditure had their share reduced to the benefit of the poorer states located in the North and Northeast.

In Chile, the government created a Municipal Common Fund (MCF), which is a horizontal equalization fund that receives up to 60% of the wealthier municipalities’ own-source income.

---

50 Baeza and Cabezas (1999) provide a throughout analysis of the experience with prospective funding and risk-adjusters in Latin America.

51 Source: Musgrove 1996 HRO Series, World Bank, Washington, DC.
and redistributes it to the poorer municipalities on the basis of a per capita formula. The formula takes into consideration the percentage of people living in rural areas, as well as the municipalities’ capacity to generate their own revenue. As table 8 shows, the fund has significantly contributed to decreasing variability in per capita funding across different municipalities (the ratio of total public health expenditure per capita in the wealthiest to the poorest municipalities was reduced from 2.2 in 1991 to 1.6 in 1996 [Bossert 2000]).

<table>
<thead>
<tr>
<th>Table 8: The Impact of the Municipal Common Fund in Chile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Income before MCF</td>
</tr>
<tr>
<td>Average</td>
</tr>
<tr>
<td>Variation coefficient</td>
</tr>
<tr>
<td>Gini coefficient</td>
</tr>
</tbody>
</table>

Source: Bossert, 2000b

Chile also provides an interesting example of how the capitation formula can be complemented by other funding mechanisms to support specific initiatives at the local level and to achieve specific pro-poor impact. The government utilized Matching Grants, piloted in the District of Nuble (Bossert, 2000b), to channel local funds towards government priorities. In Nuble the Matching Grant instrument was used to finance a program intended to fight poverty in the area by providing new services of public utility, education, as well as health services. The central government covered about 60 percent of the total costs of the project, and the local government the rest 40 percent. Bossert (2000b) reports that the impact on the poor was positive.

Resource allocation reforms have also been undertaken in South Africa, Uganda, and Zambia. In these countries, the implementation of the new resource allocation formula was not without difficulties, and progress has been uneven. In Zambia a new population-based formula was implemented in 1994, with some crude adjustments to reflect cost and need indicators. The absence of a fuel station or a bank was used as a proxy for underdevelopment, and districts received an extra weight where these facilities were missing. Subsequent research (Lake, Mtonga, and Nakamba 2002) showed that the correlation between these first crude need indicators and more sophisticated measures of relative deprivation (derived through principal components analysis) is indeed extremely significant, and that other simple indicators such as population density and remoteness can explain most of the provinces’ relative deprivation. One key issue that hindered the impact of the new population and need-based formula in Zambia was that a large share of the total public health expenditure was “top-sliced” to fund tertiary care facilities, concentrated mainly in Lusaka.

In South Africa, which had before 1994 one of the more inequitable health systems in terms of resource allocation (McIntyre and others 2000), in the immediate post-Apartheid period there...
was a strong political commitment in favor of a radical shift of resources away from the
relatively prosperous areas and toward the poorer provinces, mainly in the north and eastern
parts of the country. The allocation to the poorest Northern province was expected to increase
from between 6 and 7 percent of the total share to 15 percent over a five-year period, while funds
for Gauteng province were to decline from 25 percent to between 17 and 18 percent, a significant
reduction in real terms. Moreover, 30 percent of this shift was expected to occur in the first year.
Not surprisingly the above targets proved impossible to achieve; Gauteng received extra support
in the short term and Northern province was unable to absorb the additional funds effectively.
Also, the devolution of powers to the provincial governments (which started to receive a block
grant without specific allocation to health) and the new macroeconomic policies hampered the
redistribution process (Gilson and others 1999, McIntyre and Gilson, 2002).

This brief presentation of the RA reforms indicates that the new population and need-based RA
funding formula can achieve a strong, positive impact on geographical equity (Bossert, 2000 and
Pearson 2002). It also suggests that there is a great potential for further expanding population
and need-based RA formula, and for using more articulated risk-adjusters to pursue equity
priorities. As a first step to develop RA formula, up-to-date demographic census data would
suffice. In fact, the experience in Zambia (Lake et al., 2002) shows that significant progress can
be made simply with reasonable census data on population and age/sex breakdowns, plus some
proxy indicators of deprivation. Then, to fully develop a need-based RA formula would require
information on the following variables:

1. Standardized Mortality Rates;
2. Morbidity measures;
3. Share of population in poverty;
4. The amount of other funding sources, for instance of donors’ expenditure in each sub
   national unit.

Data on the above variables are at least partially available in middle-income countries, but are
rarely, if ever, present in low income countries. Thus, more sophisticated formulae can be more
sensitive to health needs, but they are too demanding for low income countries.

The existing experiences suggest that a number of preconditions must be met in order to support
the implementation process of the new population and need based RA criteria:

- Government’s funding must be a non-negligible part of total health funding, and the
government can decide how to allocate geographically.
- Supporting changes and improvements in overall planning procedures need to take place,
  particularly those related to personnel and capital planning and management. Adequate
  information systems must be developed to guarantee access to timely and reliable data.

53 It is important that all the variables used in the RA formula cannot be influenced by the level of government
receiving the funding.
54 RA formulae can and should then be updated as necessary and when new data become available.
• There must be basic capacity to regulate financial flows - i.e. to ensure the resources get to the intended destination, and reach the intended beneficiaries.

Many of the above conditions may not be feasible in the poorest, capacity constrained countries, characterized by $3-4 per capita public expenditure for health, and 80-90 percent of total resources for health care spent on an out-of-pocket basis. For example, consider the situation in several of the countries of Sub-Saharan Africa, where actual releases bear little relation to budgets, flow of funds is erratic, doctors, nurses and other key ‘productive factors’ are in fact engaged in different activities and away from where the financial flows records suggest, and most of the resources going to the health sector are from patients directs payments. In such a context new RA mechanisms would not be able to steer incentives towards more effective and equitable objectives.

The Priority Setting Component of RAP Reforms

Several countries have started to define an essential health package (BBP, Basic Benefit Package, or ESP, Essential Services’ Package) heavily influenced by the methodologies set out in the 1993 World Development Report (WDR), to be used as a guidance for priority setting in public funding and provision. However, in the majority of cases the implementation of the BBP has never passed the preparatory phase, or it resulted in some tokenistic change under international donors’ pressure56.

The move towards more cost-effective or more pro-poor services has been sometimes implemented at the local level. In other case, such as Uganda, reprioritization policies defined at the central level have clashed with the concomitant decentralization process57.

The few existing experiences where BBP-based policies have been really put in place indicate mixed results. In several cases essential packages have proven to be vaguely defined, setting out a range of services but giving little detail on who is to provide them, how they are to be delivered, how many services are to be delivered and to whom. For example, in 1995 the Government of Georgia introduced a new rationing policy inspired by the BBP. However, the

55 The RA formulas should be reflected in a ‘medium term expenditure framework’ and in national infrastructure development strategies, where they exist, and in long term planning documents.

56 Bobadilla (1996) presents a review of several National preparatory studies for the definition of an essential package of services, based on the analysis of the burden of disease and on the availability of cost-effective interventions. He reports that very few of those studies became translated into concrete proposals for reprioritization of government services, and even fewer of these proposals were actually implemented.

57 In the case of Uganda, for example, the amount of public funds spent on primary health care at the district level declined dramatically following decentralization despite it being a clearly stated national priority. The response there was to shift back towards a direct allocation approach in the short term by earmarking funds to the health sector. This caused problems, as some saw it as inappropriate decentralization. They thought that the new RA processes, by allocating a block grant to the local level, was supporting capacity building and more effective management at district level. Others saw it as a legitimate move by Government to ensure national priorities were followed. Recently in Uganda the process of decentralization has been moving ahead again. The role of district hospitals in delivering primary health care services has now been recognized and attempts are being made to allow these units to qualify for Poverty Action Fund financing.
range of services included in the BBP (originally only 6 programs) was quickly expanded under political pressure. The criteria for inclusion in the list of beneficiaries in the program for vulnerable groups and other programs also became extremely arbitrary. The Government plans were never enforced, and because of the fiscal collapse of the country informal payments became the prevalent source of funding for health services, both for BBP and non-BBP services.

The content of an essential service package is obviously extremely important. Ensuring equal access to an inappropriate range of services does not get us very far – a pro poor policy requires a greater focus on those diseases, which disproportionately affect the poor, and on those services that are likely to lead to the greatest health improvements for the poor. The two principles which should guide the priority setting process (the process of deciding what is ‘essential’ and what is ‘non essential’ or ‘discretionary’) should be: a) correcting market failure; and b) improving the pro poor impact of health policies (see Annex 3). By contrast, although many of the interventions outlined in the WDR 1993 basic package do address the health needs of the poor, this is by accident rather than design, as the criteria for identifying the interventions included is cost effectiveness not equity. The cost-effectiveness criterion cannot be used by itself to guide reprioritization policies (see Annex 3).

Should priorities be defined as packages of services across levels of care or rather as particular levels of care? The theoretical discussion seems to have followed the first approach, but in fact, where reprioritization policies have concretely been implemented, they have mainly consisted in the redistribution of resources across levels of care (away from tertiary and towards primary care). Prioritization policies have been more successful when they were based on existing patterns of expenditure, and proposed realistic changes at the margin. It has proved exceedingly difficult to practically disentangle the resources going to essential vis-à-vis non-essential services within primary care or outpatient care services. In Bangladesh, for example, the Essential Service Package has been defined and measured as: ‘all primary care interventions delivered at thana (district) levels and below’.

Finally, the experience from industrialized countries (Medicaid scheme in Oregon, USA, Sweden, UK, etc.), where resource prioritization and explicit rationing policies were implemented, indicates that exclusively expert-driven, radical reprioritization processes are unlikely to be accepted and to succeed. Articulating new rationing or reprioritization policies is a complex political process that is likely to succeed only if it done gradually, trying to involve major stakeholders from the outset.

---

58 As we saw in the section on equity, when the focus is shifted from services to people, cost-effectiveness and equity motivated reprioritization policies do not lead to equal results, because those that are (marginally) more able to benefit (and should then come first according to cost-effectiveness) are generally the wealthier and/or those already in better health (see Table 5 and discussion thereafter).

59 Such definition excludes ESP services being delivered in hospitals, and it assumes that all spending at thana and below is on essential services’ (Bangladesh Public Expenditure Review 1999/2000).

60 In West European countries the long debate on rationing or reprioritization in health has so far led to the exclusion of only relatively marginal services, such as largely ineffective treatments or high frequency, low cost interventions. More significant results and savings have been achieved only in containing pharmaceutical
Purchasing/contracting reforms

Let us start from the term “purchasing”, and by defining what a health purchaser or commissioner is. The purchaser is an agency interposed between patients and providers, whose role is to “organize specific types of health care for a designated population (whether defined by geography, employment type or voluntary enrolment)” (Rice and Smith, 2000, p. 1). The purchaser can be a local government, a local administrative board, a group of primary care doctors, a sickness fund or a private insurer.

The motivation of the purchasing role can be explained as follows: it is possible to have a more quality-and also cost sensitive demand, and still not renounce universal insurance coverage where such coverage is already in place, by creating purchasers charged with the role of screening demand, and of bargaining with providers on patients' behalf. These purchasers are likely to be better informed, and to have a greater bargaining power than patients do. Moreover, their purchasing power, unlike patients', can be easily equalized.

The interposition of a purchasing agency between patients and providers is a relatively new phenomenon. Until two decades ago, in every health system resource allocation decisions were organized either from the central government, as in National Health Systems, or directly through private patients-providers interaction, with little control over them by third-party insurers.61 “Purchasing” reforms were introduced first in some industrialized countries, such as the UK (1990), and then during the ‘90s in several middle income countries of Latin America (Colombia), and of the former socialist economies (FSE) in East Europe (Czech republic, Poland, Baltic Countries) and in the Caucasus (Georgia). All of these countries entrusted newly formed agencies (such as social insurance funds), or existing ones (such as local authorities, or private insurers in Latin America) with the purchasing role.

Reformers claimed that the new health purchasing agencies, separated from providers of services, would provide renovated pressure and incentives for cost-containment, efficiency and/or quality enhancement through:

1. Contractual specification of mutual responsibilities with providers.
2. Stimulating competition (for the market or within the market) among providers.

However, some governments have been reluctant to devolve the more important purchasing responsibilities (choice of providers and of payment systems) to decentralized purchasing agencies, and, where they have been created, the latter have actually assumed more the role of cashiers of the delivery system, rather that the funding/purchasing role proper. In other countries expenditure, by defining different categories of drugs subject to different reimbursement policies. Part of the burden of financing pharmaceutical expenditure has thus been shifted on to households.

61 For instance, neither the British nor the American health systems, which exemplified the two opposite mechanisms for allocating resources among industrialized countries, had agencies entrusted with the purchasing role. The British National Health System was, at least formally, a centralized and vertically integrated system where Health Authorities, the public agencies locally responsible for planning and provision of health services, directly managed hospitals. In the American health system, on the contrary, insurers reimbursed health providers mainly on a retrospective, fee-for-service basis (see discussion on payment systems in the text).
reforms did decentralize purchasing decisions to agencies separated from the government, and real competition for patients was introduced across purchasers. In the next section we discuss a possible trade-off between equity and efficiency that this new competition between purchasers may generate.

Other countries, without relying on purchasing agencies separate from the government, did nonetheless experiment contracts to explicit discipline the relationship with private not-for-profit or for profit providers. The three forms of contracts more widely utilized have been the following (Loevinsohn, 2001):

- **Contracting Out (CO):** in which the contractors assumes complete responsibility for service delivery, including hiring, firing and setting wages, procuring and distributing essential drugs and supplies.
- **Contracting In (CI):** where the external contractors is brought in to provide certain specific productions inputs, but always within the umbrella of the public health system, managed by the MOH. NGO may for example be asked to manage the health centers in a particular district, using the logistics, drugs and supplies provided by the government. The contractors cannot hire or fire health workers, and any surplus can be used to provide incentives for staff, or to cover operating expenses, etc.
- **“Leasing”** where an organization offers to manage and finance existing health facilities (e.g. an NGO offers to take over and manage poorly performing health centers, provide additional financing, and upgrade the existing staff).

Still very few rigorous studies exist to assess the overall performance of the new purchasing and/or contracting arrangements, particularly its impact on equity, and most of the evidence available is from gray literature, from retrospective studies or from relatively small-scale projects. In the section on the trade-offs, we highlight a possible trade-off between equity and efficiency which may become evident as the new contracting experiences are scaled-up.

In synthesis, the evidence collected so far reveals both promising as well as problematic aspects of the new contracting experiences. The evidence from Cambodia (Bhushan and others, 2002) suggests that contracting in and out services to NGOs produced an extremely positive impact on the poor, improving coverage and quality of services for lower-socio-economic groups, as the following Table 9 indicates. The benefits were greater for contracted out districts, where government service delivery was more expensive ($4.50 per capita, compared with $ 2.82 and $1.5 spent per capita in contracted in and control districts), but where private out-of pocket payments were reduced by 27 percent over all and over 70 percent for the lower socio-economic groups, indicating good targeting of beneficiaries.
Table 9. Average Change in Service Coverage Indicators in Cambodia (first 2 and ½ years of reforms)\(^6^2\)

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Control District</th>
<th>Contracted-in</th>
<th>Contracted out</th>
</tr>
</thead>
<tbody>
<tr>
<td>Immunization rates-all(^6^3)</td>
<td>56</td>
<td>82</td>
<td>158</td>
</tr>
<tr>
<td>Vitamin A capsule Receipt –all</td>
<td>-25.1</td>
<td>18.1</td>
<td>20.9</td>
</tr>
<tr>
<td>Vitamin A capsule Receipt –lower 50% socio-economic segment of population</td>
<td>-24.1</td>
<td>29.9</td>
<td>23.9</td>
</tr>
<tr>
<td>Percent of illnesses Treated in Public Health Facility—lower 50% socio-economic segment of population</td>
<td>81.7</td>
<td>490.5</td>
<td>1096.0</td>
</tr>
</tbody>
</table>

Source: Bhushan and others (2002)

Utilization of curative services in district hospitals by the bottom 50 percent socio-economic group increased about twelve fold in contracted out districts, six fold in contracted-in districts and only less than double.

The same positive results are reported in local studies which analyzed contracting with not-for-profit providers (NFPP) in Bolivia, Bangladesh (nutrition services), India (TB-DOTs treatment), and sub-Saharan Africa (Giusti, 2000)\(^6^4\). The flexibility in the management of inputs allowed NGOs to improve service coverage, and achieve significant cost savings compared to government standards without compromising quality of services (for a careful review, see Loevinsohn and Harding, 2004). A study of the impact of contracts in South Africa (Palmer 2001) indicated the importance of maintaining competition among providers, and of defining ex-ante the population to be covered by the contract. Fixing the beneficiaries’ pool allows to use capitation payments, which encourages prevention.

\(^6^2\) In Cambodia two forms of contracts were utilized in different districts (Loevinsohn, 2001):
- Contracting Out (CO): in which the contractors assumed complete responsibility for service delivery, including hiring, firing and setting wages, procuring and distributing essential drugs and supplies.
- Contracting In (CI): where the external contractors was brought in to provide certain specific productions inputs, but always within the umbrella of the public health system, managed by the MOH. NGO were asked to manage the health centers in particular districts, using the logistics, drugs and supplies provided by the government. The contractors could not hire or fire health workers.

\(^6^3\) Baseline immunization coverage was 25.5 percent in contracted out districts (C0D), 29.9 in contracted in districts (CID), and 34 percent in the control districts (CD). The follow-up survey showed coverage of, respectively, 65.8 (COD), 54.4 (CID), and 53 (CD).

\(^6^4\) In Uganda, the case analysed by Giusti (2000), the government contract and subsidy allowed NFP hospitals to reduce user-fees, and this created a positive impact on utilization.
Several other authors have pointed at the significant human capital and information systems requirements to specify and to manage contracts effectively, to be contrasted with the limitations characterizing several developing countries (McPake & Banda, 1994, Mills et al., 1997, Palmer, 2001, Taylor, 2000, Van Holten, 1998, Vining and Globerman, 1999, Waters et al. 2002). From the evidence available, one can say that effectively designing and monitoring contracts has proved as demanding as directly managing health services, and that where the public sector was not able to effectively ‘row’, it has generally also proved unable to effectively ‘steer’ independent or semi-independent providers through contracts. Thus, lack of funds and managerial capacity, absence of competition, and in some cases widespread corruption blighted the impact of contracts. The following case study from Bangladesh illustrates some of the problems encountered in contract management.

**Case study 5: Bangladesh experience with engaging the not-for profit sector in delivering nutritional services**

In Bangladesh, there is an ongoing experience with contracting with NGOs for nutritional and primary care services in two large projects, the Bangladesh Integrated Nutrition Project (BINP) and the Urban Primary Health Care Project (UPHCP). BINP, for example, has contracted with 7 NGOs to provide nutrition related services to a population of about 10 million and has been in operation since 1997. Loevinsohn (2001, p.1 and subsequent pages) reports: “Contract management has been a significant issue under BINP with one NGO owed more than $3 million and the other contractors also owed substantial amounts. There have also been significant problems with the Government’s procurement and distribution of supplies and equipment, particularly in comparison with those materials procured by the NGOs, which have been readily available. Field supervision and monitoring by BINP has been problematic. The mid-term review of BINP examined the change in nutritional status in 6 project thanas and 2 control thanas as judged by a third party. It appears that the decline in malnutrition in project thanas was, at best, only modestly greater than in non-project thanas”.

One example of the use of new purchasing mechanisms to improve services for the poor is the experience of contracts with not-for-profit providers (NFPP) in sub-Saharan Africa, where NFPP manage a significant share of the health services serving the rural and poorer segments of the population. In particular, in several countries facilities run by religious organizations have traditionally been closer to service beneficiaries than government facilities. In the past, NFPP were completely separated from the government, and were entirely relying upon user-fees and private donations. In recent years, several governments have begun to acknowledge their role, offering subsidies in exchange for agreements on service provision and price controls (for example in Uganda). Sometimes targeted beneficiaries of the services are poor and marginalized groups. The government subsidies allow the NFP sector providers to decrease prices charged to patients, thus enhancing accessibility of services for the poor. Again, we must underline the lack of rigorous evaluation studies on the impact of such new experiences. However, the existing qualitative evidence, suggests that at least in a few cases the new arrangements for service delivery have been able to significantly increase health care utilization by poor and marginalized groups, particularly when communities and users have been involved in the planning and the monitoring of services.

---


66 According to a document reviewing the experiences with contracting with NGOs in Western Africa: “The factors that favor this community involvement include “well-organized NGOs; political willingness; and democracy. With
Payment systems’ reform

Payment systems play a particularly important role in the health sector, due to the peculiar interplay of demand and supply in health. Probably more than in any other market, when demanding health services individuals are often largely uncertain about the appropriateness of the service they are receiving. Any health service is an experience good, a good whose quality can be ascertained only ex-post. Even ex-post, the quality of any individual health service is difficult to precisely assess because the patient often receives a multidimensional service (diagnosis, advice, treatment...), and because the final outcome, ‘good health’ or ‘health improvement’, which is what people really demand, depends on factors other than the kind of care received, the output produced. The lack of information and the particular psychological condition of patients produce crucial consequences for the interaction between demand and supply. It is the doctor, and not the patient, who generally chooses the amount and the type of services that the patient needs.

Hence, market outcomes in health are largely determined by the incentives, as well as the cultural influences and professional ethics on the basis of which providers act, without strong influence from the demand side (more on this in the section on trade-offs). Payment and reimbursement criteria, in turn, are critical to determining those incentives, and that is why different payment systems can lead to completely diverse outcomes, in terms of accessibility, quality and cost of services offered within the health sector.

Several countries of Latin America, Asia, and Africa implemented radical payment systems’ reform over the last decade. The general trend has been to move away from input-based payment systems, towards output-based payment systems, such as DRG-based payments for hospitals and fee-for-service reimbursement for primary care, and, more recently, towards fixed-prospective budgets for hospitals and capitation payments for primary care doctors (see Box 1). With the new resource allocation criteria introduced through RAP reforms, part of the financial risks for unexpected increases in unit costs and volumes of service is shifted on from the central government to local governments or health purchasers, and, in turn, the latter ones shift part of these risks on to providers through prospective payment systems.

We present the major payment systems in the following Box 1 and the evidence available from PS reforms in Colombia in Case Study 6. Unfortunately, for all countries which undertook PS reforms we do not have data which could be used for an equity assessment (such as changes over time in utilization of services by socio-economic group), but only data showing the aggregate evolution of hospital activity indicators after the reforms. The evidence from Colombia, Hungary, Poland, and the several other countries which moved from input towards output and

regard to NGOs, their participation is encouraged because of their efficacy, their flexibility in their interventions, their ability to target beneficiaries and interventions, their innovative approaches, and their cost-efficiency (From Partnership to Contracting: Report of the inter-country workshop on contracting private providers on health services, Bamako, Mali, June 12-15, 2000, p.1)\(^\text{67}\).

\(^{67}\) Similar, although less severe, problems arising from the asymmetry of information between supply and demand occur in other sectors, such as the financial investment or legal advice markets.
case-based payment systems for hospital care consistently indicates increased activity volumes and often improved productivity indicators (decrease in length of stay and unit costs).

We will take up again some of the key issues regarding the evidence from payment system reforms in our conceptual discussion of the trade-offs (last section). We will argue that theory indicates that where the purchasing and payment systems’ reform components were devised mainly to enhance efficiency or reduce costs, they may actually cause an adverse impact on equity.

**Box 1: Different payment systems for health providers**

First, it is important to distinguish between reimbursement systems for individual providers and those for provider facilities. Usually, the criteria for paying individual physicians and medical personnel are influenced by those for their facility (hospital, clinic or health centre), but this is not necessarily always the case. For example, in several countries of Western Europe physicians continue to receive a fixed salary independent of their output, as they did a decade ago, even though over the last decade the reimbursement criteria for provider facilities, such as hospitals, have radically changed, from input towards output-based payment systems.

Focusing first on facilities, until two decades ago across OECD and developing countries one could find two major reimbursement systems:

- In public health systems provider facilities were mainly reimbursed according to inputs and historical expenditure. Financing was distributed separately for the different line items of the provider’s budget (so much for salaries, so much for capital, so much for consumables, etc.), and the total budget was generally rolled over from year to year with minor adjustments. At the same time, health workers received a fixed salary generally dependent only on seniority, and they were all included in the civil service. Under such payment system, the government would ultimately bear any financial risk for any unforeseen increase in unit costs or volumes. Although an input-based payment system does not provide incentives for efficiency, or cost-control, at the macro level health expenditure was governed through a rigid planning of price, quantity and mix of inputs. This form of cost control was relatively effective, but, according to many critics, did not provide any incentives to improve services (see previous section). The total amount spent and the share attributed to different services continued to be determined through a political and bureaucratic process in a way that hardly came to reflect individuals’ preferences.

- In parallel, in the private sector providers were mainly paid retrospectively (that is, with full cost-reimbursement), on a fee-for-service basis, directly by patients and/or by indemnity insurers. In a fee-for-service payment system the revenue facilities (and individual doctors) receive depends on the volume and the value of the services they provide. Thus, those OECD countries, such as the US, where this kind of reimbursement system was more widespread, over time experienced a pronounced health expenditure growth, in absolute terms as well as a proportion of GDP. One stream in the US health economics literature underlined that, when reimbursement to providers is activity-and-cost-based (as in a fee-for-service system), there is no incentive to focus on technological progress that could lead to less costly treatments. On the contrary, providers can gain by making use of ever more costly treatments and equipment, and even by inducing demand and supplying services above the level that would be clinically justified. This phenomenon, named ‘supply-induced-demand’ (SID), has been extensively studied in the health economics literature.

SID was first studied by Evans (1974). Later, Fuchs (1978) emphasized the welfare-reducing shift in the demand curve induced by providers through advertisement, and Robinson and Luft (1985), Dranove (1988), and Mooney...
In the last twenty years, in all OECD countries and some developing countries new and more sophisticated payment systems have been introduced.

- Case-based (or fixed-price, according to US terminology) payment systems, first introduced in the US Medicare system (1983) and later on adopted in several other countries, link payment to hospitals to the number and the severity of the cases they treat. Each patient is classified in a specific “diagnostic” group according to his/her principal diagnosis and, correspondingly, a fixed reimbursement is given to the hospital for treating the patient. When the system was initiated, all pathologies were classified in 470 different Diagnostic Related Groups (DRG); since 1983 the number of DRG categories has been expanded. In the late ’80s and ’90s, case-based payment systems were adopted in several developing and former socialist economies (FSE) to pay for inpatient care (the first government to adopt them in Latin America was Brazil and among FSE was Hungary), in some cases with the same classification system (DRG) as in the US, in other cases with other classification systems (for example, the “Nosology-based” system in several of the Former Soviet Union countries). Simpler classification systems are easier to manage, easier to control against manipulations and provide stronger incentives for cost-control. On the other hand, in simpler classification systems set prices are less likely to reflect real relative treatment costs, they may lead to budgetary unbalances for providers and create the incentive to focus on more profitable cases/diseases. As the disease/treatment classification system becomes more articulated its effects approximate those of a fee-for-service payment system. The opportunities of manipulation (the so called DRG creep) decrease with the incentive to contain costs.

With a cost-and-volume contract, similarly to a cost-per-case payment, providers receive additional funding for each extra case they treat, but additional units of service are reimbursed at a progressively lower unit price.

Capitation payment for primary care is a lump sum payment for each patient enrolled with a particular physician, which can be adjusted in relation to age, sex and past medical record of each patient. Two variants are used to determine enrolment: in the first, each patient resident within a specific geographical area is automatically enrolled with the same physician; in the second, each patient can choose his/her primary care physician, also among those practicing outside his/her area of residence. Capitation payment and Ryan (1993) gave a theoretical explanation of supply-induced demand, by pointing at the imperfections in the agency relationship between physicians and patients. The phenomenon of supply-induced demand has also been studied by a vast empirical literature (starting from Roemer, 1961), which found a strong correlation between supply and utilization of health services (by examining cross-the sectional correlation between bed-supply and utilization, Roemer found that "a bed built is (roughly) a bed half-filled"), and a positive association between physicians’ ownership of testing facilities and the number of referrals for diagnostic tests (Crane, 1992, and Mitchell and Scott, 1992, studied the relationship between the number of test prescriptions and physicians’ ownership of imaging equipment, and found a strong positive correlation between the two). Despite this empirical literature, it must be said that unequivocal evidence of demand inducement has not been found, and that most of the literature in support of the supply-induced demand hypothesis has been questioned for its econometric shortcomings. Dranove and Wehner (1994), for example, by using TSLS regression analysis, found evidence in favor of supply-inducement of child-births (!). This lead them to question the identification of the demand inducement econometric models, as they found that the “exogenous” instrumental variables used in the first stage of the TSLS estimation, which are the same as those traditionally used in most empirical works aimed at estimating supply-induced-demand, in fact are not exogenous. Thus, despite a good economic understanding of the rationale behind the supply-induced demand phenomenon, its empirical relevance is still controversial.

69 Presently, DRG codes are based on the 9th revision of the International Classification of diseases.
provides incentives to avoid unnecessary treatment. However, physicians may also be induced to under-treat and to risk-select patients (see discussion of second trade-off).

A fixed-budget (or block contract) is a lump sum prospective payment assigned to health facilities for the provision of a pre-specified volume of services over a certain period of time. It maximizes incentives to control costs and to avoid unnecessary treatment, and it is administratively simple to manage. However, it may lead to under-provision and to selection of patients, unless explicit funding is provided for quality enhancement and for treatment of extra-expensive patients (see second trade-off).

Fee-for-service and input-based payments are frequently referred to as retrospective payment systems, because the cost of services is completely reimbursed ex-post. On the opposite side of the spectrum, block contracts are referred to as a purely prospective payment system, because the amount of funding the provider receives is determined ex-ante, and completely independent of the number and the cost of the cases treated. Sometimes, erroneously, cost-per-case or fixed price payment systems are also referred to as prospective payments to contrast them with the fee-for-service arrangements they replaced. Moving from fee-for-service, towards DRG-based reimbursement, cost-and volume contracts, capitation funding and block contracts progressively increases the part of providers” payment that is fixed ex-ante, thus increasing also the financial risk associated with unexpected increases in the volume or and the cost of services.

Case study 6: Hospital autonomy in Colombia (Extracted from McPake et al. 2002)

Before the reforms in Colombia efficiency indicators in the public sector were extremely poor. For example, occupancy rates in public facilities were lower than 40%. As explained before, the reform programme included the creation of a purchaser-provider split and the transformation of public hospitals into ‘autonomous state entities’. These were intended to contract with multiple competing insurers and the local health secretariat for the provision of services. A study aimed to track hospital performance in the post-reform period in Bogotá (McPake 2002). Trends in hospital inputs, production and productivity, technical quality, patient satisfaction and finances, and qualitative data based on interviews with hospital workers were collected. There was some evidence of increased activity and productivity and sustained quality despite declining staffing levels. The following two figures show trends, respectively in total admissions and in bed occupancy rates.

Figure 8: Trends in Total Admissions

Source: McPake 2002
As the two figures above show, both number of admissions and bed-occupancy rates either remained stable or raised after the reforms, as one would expect after the introduction of a case-based reimbursement system which replaces an input-based payment system. Qualitative data suggest that hospital workers have noticed considerable changes, which include greater responsiveness to patients but also a heavier administrative burden. Unfortunately, no data is presented concerning the variation in service utilization by the poor.

The Use of RAP Arrangements in Specific Pro-poor Health Interventions

Supply-side incentives

To achieve better coverage and improved health outcomes among the poor some countries have introduced explicit performance-related financial incentives for providers, particularly in their vertical programs against communicable diseases. This approach has also been adopted by GAVI, the new global vaccine initiative, where an amount equal to US$ 20 is assigned for each additional child immunized (DP3) in countries with immunization coverage rates comprised between 50 and 80 percent. Similar ‘pool’ (as opposed to “push”) mechanisms are being considered for the Global Fund disbursements, to fight malaria, TB and HIV/AIDS.

We believe that these explicit financial incentives to treat the poor/diseases that principally affect the poor have a great potential, and they could be used more extensively. They could also be used to correct the potential negative impact of other revenue collection measures, such as increases in user fees in government facilities, meant to raise additional revenue. Consider for example the introduction of dual-charging (DCM) mechanisms in public hospitals. Under the DCM, poorer patients maintain the option of using services for free in wards or beds reserved for non-paying patients, whilst wealthier patients can seek care in the other wards. There is evidence that suggest that such dual charging mechanism lead to adverse equity consequences, widening disparities in quality of treatment (Nakamba, Hanson and McPake. 2002). Again, such adverse effects can be contained by paying a premium to providers linked to the number of patients they admit in the non-paying wards. In Case study 5, we review the experience on the use of enabler and incentive schemes to improve coverage and outcome of TB treatment, based on Direct Observed Treatment (Weil et al., 2001).
Case study 7: The Use of Patient and Provider Incentives to Improve TB Program Performance

Tuberculosis (TB) remains a major killer of adults worldwide, only surpassed by HIV/AIDS, with two million persons dying due to TB last year and the toll worsening rapidly in some regions, especially Sub-Saharan Africa with the intimate linking of the two epidemics. The poor are hardest hit by TB, given the nature of transmission of infection and breakdown to disease, but all income strata are at risk. A proven cost-effective strategy to detect and cure TB control that is capable of reducing mortality and morbidity even in very poor settings is known as Directly Observed Treatment Strategy (DOTS). Among the top challenges faced by DOTS programs throughout the world is ensuring that patients are able to initiate treatment quickly, that they stay motivated to complete the full course of therapy, and that providers assist patients in this process as well as provide safe and appropriate care.

Recognizing the need to improve DOTS-based TB control programs, several projects have introduced schemes that provide incentives to enable and motivate providers and/or patients to pursue DOTS functions and objectives. Some incentives are not linked explicitly to performance, but others relate directly to measures of improved performance. Examples of the former include providing improved training opportunities for TB workers or per diems for service supervision; examples of the latter include providing food packages to patients that present regularly to take their medicine and providing financial bonuses to providers for each infectious patient that completes treatment. Some of the most successful DOTS programs worldwide, such as in China, Peru and Cambodia, use performance-linked enablers and incentives.

A preliminary survey of the literature on enabler and incentive schemes linked to TB control projects showed that few schemes involved rigorous design, monitoring or evaluation elements, or had clear financing and management plans (plus, among the existing 23 publications only 2 were from developing countries). No studies focused on incentives in contractual relationships between governments and private or non-profit providers; and no studies attempt to measure the cost-effectiveness of schemes. The survey of the literature also showed that design, implementation, and impact of incentives schemes for TB treatment vary significantly. Nonetheless, the limited developing country evidence combined with the more extensive developed country evidence indicates that incentives for providers and patients influence behavior and have the potential improve program performance.

The survey of the existing experiences also showed that there was considerably more concern about formal provider schemes than patient-focused schemes. This apparently related to the strong perceived potential unintended negative effects on performance in other areas or for future motivation if resources for incentives are no longer available, but more evidence is needed. Some of the more significant conclusions reached so far are summarized in the following:

i. Problems in motivating and enabling patients and providers are important and relevant to effective TB control throughout the world and may inhibit DOTS expansion and achievement of TB control targets.

ii. These problems may especially inhibit access and utilization by poor patients or participation by providers working under difficult conditions.

iii. Incentives exist for all stakeholders in TB control – whether they are explicit or implicit. Some of these incentives may inhibit good performance and are not aligned with the objectives of TB control.

iv. Even with DOTS implementation, some functional constraints inhibit patients and providers from participating fully and effectively. Formal incentives that encourage actions to be in line with DOTS norms may improve performance.

v. Enablers and incentives may be appropriate and help improve coverage and cure rates in DOTS programs, which have already achieved cure rates of 70% or more. They may not be right to consider if

---

70 The case study is an extract from work by A. Beith, R. Eichler, J. Sanderson, D. Weil.

71 Some schemes are clearly defined and tied to measurable actions or results while others are less regularized. Some schemes depend on sporadic funding while others are fully incorporated into their program’s annual budget. Some are offered by the public sector and others through NGO or private providers.
fundamental problems of standards, inputs, training and supervision have not yet been at least partially alleviated.

vi. Appropriate enablers and incentives may vary considerably depending on underlying epidemiological, economic and social conditions, health system structures and capacity. There is likely a continuum of approaches that may be effective from specific local measures to standardized national approaches.

vii. There are potentially serious unintended negative effects of enablers and incentives that must be avoided through proper design and monitoring.

**Demand side incentives**

Explicit demand-side incentives can and have been utilized to stimulate consumption of specific health and reproductive health services. Demand for health services is usually lower among the populations needing them most. To stimulate the consumption of services with large externalities (vaccinations, maternal and child care, communicable disease treatment), the opportunity cost (that includes time lost for gainful employment, traveling costs, etc.) can be lowered for the poor, by providing specific enablers or incentives (such as vouchers that contribute to treatment and traveling expenses). These demand side RAP mechanisms are investigated, among others, by Gorter et al. 1999 and by Armstrong et al., 2002. Gortler et al. report the experience of vouchers for sex health services in Nicaragua, and Armstrong et al. report the experience with vouchers for mosquito nest in Tanzania. Both programs were not without difficulties, but they overall lead to significant increases in utilization and improved health outcomes. Case study 6 reviews the experience in Nicaragua.

**Case studies 8. Vouchers for sexual health services in Nicaragua (Source: Gorter et al. 1999)**

Since 1995 a voucher programme has been operating in Managua, Nicaragua, with the aim of increasing the uptake of reproductive health services by female sex workers. Every 3-5 months approximately 1200 vouchers (corresponding to the estimated number of sex workers operating at any given time in the city) are distributed by fieldworkers and NGOs at prostitution sites. The vouchers entitle the sex workers to free services at one of 8-10 private, NGO and public clinics, which are contracted to the voucher agency by competitive tender. Approved providers must follow a set treatment protocol, and receive training. Contracts are reviewed after each round of voucher distribution, and renewed subject to an assessment of quality of care. The clinics return the vouchers to the voucher agency, which reimburses the provider an agreed fee per voucher. The sex workers were involved in the design of the programme, and have a number of opportunities to express their preferences and complaints. In each round, 10% of recipients are interviewed about their experience. Initially, sex workers reported that the gatekeepers to care (nurses and receptionists) lacked sensitivity: training. Sensitization of this group helped to improve their attitudes towards these clients. Technical quality of care (as assessed by an examination at the project outset) was lower than expected, and training and treatment protocols were introduced. While the prevalence of STD is only slightly lower than at the beginning of the project (possibly due to a high turnover of female sex workers), incidence among women who have used vouchers more than once dropped by 65% in the first three years of the programme. Following a recommendation by the sex workers, they now receive vouchers to give to their regular partners and/or clients as well. Sex workers appreciated the fact that they could chose which clinic to attend, and made their choice on the basis of distance and friendliness. The clinics reported that their main benefit was improvement in the technical quality of their services, and that the lessons learned were applied to all of their clients. They felt that their reputation was enhanced by being contracted by a prestigious public health agency (the Central American Health Institute, ICAS)
TRADE-OFFS IN THE PURCHASING AND PAYMENT SYSTEM
COMPONENTS OF RAP REFORMS

In this section, we introduce a simple conceptual framework which can be utilized to illustrate the impact of RAP reforms, and the potential trade-off between equity and efficiency. Let us utilize a graphical illustration of the (expected) Health Pareto Frontier (HFP)\(^{72}\) for two individuals, Mr. Rich (R), and Mr. Poor (P):

![Figure 10: RAP reforms. Moving Towards the Pareto Frontier and Improving the Health of the Poor](image)

The above figure simplifies a situation where the achievable health gains are constrained by the available amount of resources, and where the choices on the use of the existing health resources may create an equity-efficiency trade-off. Each point in the graph represents the (expected) health, or health gains for the two individuals associated with a different allocation of health resources and a different incentive structure within the health system. The down-sloping Health Pareto Frontier (HPF) curve represents the set of Pareto optimal points, where it is not possible to further increase expected health gains for one of the two individuals, without decreasing them for the other. Suppose that in the initial situation the health sector is characterized by a suboptimal equilibrium (point e in the graph), where due to a distorted allocation of resources or lack of incentives the system is not attaining the maximum health gains for the two individuals. Graphically, the allocation in correspondence to point e is not Pareto optimal, being an allocation internal to the HPF.

\(^{72}\) The Pareto Frontier is the set of all Pareto optimal allocations. Each point on the health Pareto Frontier for Mr. Poor and Mr. Rich is found by fixing the level of expected health of one of the two, and by maximizing expected health for the other.
Suppose that efficiency is measured by the sum of the health gains for the two individuals, and that equity is measured by the difference in the total health enjoyed by the two individuals (egalitarian view). Starting from point $e$, suppose that a RAP reform is introduced which drives the system towards a greater level of efficiency (towards the HPF). As the graph indicates, it is possible to improve efficiency, in three ways. The first is to move from point $e$ towards the Pareto frontier in area $C$. However, moving from point $e$ in area $C$ does not lead to a Pareto improvement, because Mr. Poor is adversely affected by the change (although losing less than what Mr. Rich gains). It also leads to worsen equity outcomes according to our measure, as the health differential between the two individuals widens. The second and third ways to improve efficiency are to move from point $e$ towards areas $A$ or $B$. In fact, such movements not only enhance efficiency, but are also Pareto-improvements, as both individuals are experiencing positive health gains. However, as the figure shows, movements in $A$ or $B$ do not create the same equity impact: moving towards points in area $A$, unlike points in area $B$, will also decrease inequality in health between the two individuals.

In conclusion, conceptually we understand that the purchasing and payment system components of RAP reforms, designed mainly to address efficiency and service-quality issues, may entail negative side-effects on the equity dimension, at least according to a specific more egalitarian concept of equity. In the graph, starting from any Pareto sub-optimal situation, we can point at the existence of an equity-efficiency trade-off any time as a consequence of health reforms health outcomes improve overall but also health disparities widen. In this section, we present a conceptual analysis of the following trade-offs associated with the purchasing/contracting and the payments systems dimensions of RAP reforms:

1. In purchasing, between the aim of avoiding adverse selection by maintaining large risk-pools, and the aim of allowing greater choice for clients and consumer-responsiveness of purchasers, by stimulating competition.
2. In contracting, the trade-off between the aim of maintaining uniform standards of care everywhere, which requires central control, and the aim of enhancing technical quality and results-orientation, which require more competition and incentives at the individual level.
3. In payment systems, between the objective of making both purchasers and providers cost-conscious through prospective payments, and that of enhancing technical quality and avoid selection of patients through activity-based payment systems.

Note that this concept of efficiency is different from that of Pareto-efficiency. It is possible to increase the sum of the health enjoyed by the two individuals also by moving towards points where one of the two individuals is worsening her/his health status.

For a critical review of the different concepts and measures of equity currently used, see the section on equity in the text. In Figure 10, note that if we were only interested in the absolute health of the poor (as argued by Rawls, 1971), and not in overall disparities, movements towards points in area $B$ will be judged positively from also an equity perspective. See also Case Study 2 presented in section on equity, based on Waters (2000).
PURCHASERS: CONSUMER CHOICE AND CONSUMER-RESPONSIVENESS VERSUS NO RISK SELECTION

Under the new RAP arrangements, a trade-off arises between the aims of limiting risk selection and of pursuing positive redistribution policies through the pooling mechanisms, and those of increasing purchasers’ consumer-responsiveness and patients’ choice through competition.

Reformers would like purchasers to reflect patients’ legitimate preferences. For example, to promote through contracts ‘good’ providers and to penalize ‘bad’ ones, according to patients’ perceived quality. Regardless of all other possible forms of regulation and control, the only true guarantee that purchasers will take into account patients’ preferences in their choices is to allow patients to leave them when they are dissatisfied with their purchasing choices, and to financially penalize those purchasers that are not able to attract patients. A precondition for this to happen is that patients be free to choose among different purchasers. That is why many purchasing reform proposals in public systems (for instance, Poland and Czech Republic among former socialist countries, and Mexico, Chile and Colombia in Latin America) intended to create and safeguard at least some competition among purchasers. The claim was that without any competitive pressure on the purchaser side, the allocation of resources to providers would continue to be imposed from the top down, as it was before. Providers could collude with purchasers against patients’ interest, and unsatisfied patients would have no possibility to change, or to “vote with their feet”.

At the same time, at least in health systems committed to universal insurance, reformers also require that purchasers continue to guarantee coverage at equal standards, and that they do not select patients according to their health risk. However, the two aims, to promote good purchasing through competition and to promote equitable and efficient pooling, are intrinsically in conflict with each other.

Note that this problem cannot be solved simply by assigning the pooling and the purchasing function to two different agents. Even if resources are collected and pooled “upstream” (at the central level), prospective resource allocation and payment systems shift all or part of the financial risk “downstream”, on to purchasers and on to the lower-level agent who is paid prospectively (unless the possibility of reinsurance is contemplated, see below).

Thus, the same issues that plague private health insurance markets (see Newhouse, 1996) resurge by paying purchasers prospectively and letting them compete with each other. “Good risks” might receive incomplete coverage, whilst “bad risks” may be excluded from coverage. “Bad risks” usually coincide with the poorer and more vulnerable segments of society, and so market segmentation may lead to worse services, or outright exclusion of the old, the chronically ill, and of other vulnerable groups.

75 The theory of insurance markets with adverse selection proved that some positive level of redistribution in general characterizes second best equilibria across different risk groups. The free market equilibria may fail to achieve such cross-subsidization, because, in Rothschild and Stiglitz's (1997, p. 74) words, "One of the natural methods that insurance companies use to compete, underwriting, destroys or considerably limits what insurance markets can do". Thus, the separating equilibrium achieved by the market may imply an excessive degree of market segmentation. In other words it may be in the low risk group's interest to provide some form of subsidy in favor of the high-risk individuals as a means to achieve a degree of coverage closer to their preferred full coverage contract.
As far as this trade-off is concerned, the ‘size’ of the purchaser (measured by the number of people covered) is a critical dimension. As the size of the pooled population increases, the degree of consumer-responsiveness decreases, because competitive pressures become weaker, and because purchasers become more distant from patients. However, also the incentive for risk selection diminishes, as the purchaser can exploit substantial economies of scale in risk pooling, while cross-subsidies across different risk and socio-economic groups become easier to sustain. The larger is the risk-pool, the lower is the variability in individual income, and so the more efficient is the risk-pool, due to a very simple statistical reason. We illustrate this point in Box 2.

**Box 2: Risk-pooling and the size of the purchaser**

Risk pooling can be represented as a mechanism through which risk-adverse individuals insure against their risk of income variations, by pooling together their resources. Consider \( n \) individuals, and let each individual’s income be a stochastic variable \( y_i \), where \( i=1,\ldots,n \), characterized by a certain statistical distribution, with expected value equal to \( \mu_i \) and variance equal to \( \sigma_i^2 \). For simplicity, assume that the stochastic variables \( y_i \) are independent and that the expected value and the variance are equal for all individuals (\( \mu_i=\mu \), \( \text{var} y_i=\sigma^2 \), for each \( i \)). Consider what happens if in each period individuals pool their incomes into a mutual insurance fund, that then assigns to each one the mean of the aggregate income received, thus compensating the variations in individual incomes occurring over time. The variance of the aggregate income received by the fund is equal to \( n \) times the variance of individual incomes (the variance of the sum of independent stochastic variables is equal to the sum of the individual variances):

\[
\text{var}(y) = \text{var} \left( \sum_{i=1}^{n} y_i \right) = \sum_{i=1}^{n} \text{var}(y_i) = n \sigma^2
\]

However, after the individuals pool their income in the insurance fund, the variance of individual income is equal to the variance of mean income. This is equal to \(^{76}\):

\[
\text{var}(y_i) = \text{var} \left( \frac{\sum_{i=1}^{n} y_i}{n} \right) = n \text{var} \left( \frac{y_i}{n} \right) = n \frac{1}{n^2} \text{var}(y_i) = \frac{n}{n^2} \sigma^2 = \frac{\sigma^2}{n}
\]

Thus, by pooling their income together through a risk-pooling mechanism, individuals are able to obtain income equal to the mean income, with a variance equal to \((1/n)\) of individual variance. It is clear that as the number of people in the pool increases, the variance of the mean income shrinks approximating to a situation of certainty (Central Limit Theorem). Also, note that the “risk-pooling” function is conceptually distinct from the redistributive function that may also be achieved by pooling funds. In our example, redistribution of income across individuals may be achieved if different individuals ex-ante had different expected income. By joining the pool the rich would then trade an expected higher income with greater variability with a lower expected income with lower variability. Again, ceteris paribus increasing the size of the pool would increase the convenience of pooling.

As far as risk-selection is concerned, in health matters are made worse by the extreme concentration of risks (evidence from several countries shows that in each country approximately five/seven percent of the population is responsible for roughly two-thirds/three quarters of total health expenditure). For small purchasers, who have a disproportionate share of bad risks and who are unable to dilute them in a large pool of insurees, risk selection may become the only financially viable strategy.

---

76 In the following equation, we refer to a very simple propriety of variance, namely that: \( \text{var}(ax+b) = a^2 \text{var} x \).
In theory, as we have discussed before, the resource allocation criteria can provide a strong incentive not to risk-select, by providing appropriate rewards to purchasers who recruit and retain bad risks (risk adjusters\(^{77}\)). Insurers/purchasers that have a disproportional share of chronically ill people or poor people, for instance, should receive a premium support or subsidy from a compensatory fund. Unfortunately, however, existing risk adjusters present severe limitations, and they cannot eliminate the advantage of screening risks because they predict only a minor part of service utilization.

Another way to prevent risk-selection would be to allow small purchasers to reinsure with larger ones, but then their incentive to curb expenditure would also be diluted accordingly. The point is that in order to limit risk-selection, it is necessary to pay a price either in terms of less competition, less consumer-responsiveness, or in terms of less tight incentives for cost control. Each country must find its own “optimal point” along the lines of such trade-off. We briefly discuss in the Case Study that follows the reform debate during the ‘90s in the US and the UK in light of the trade-off just explained.

**Case study 9: Purchasers in the UK and in the US**

We can represent in a graph the point chosen by the different purchasing reforms proposed in the UK and the USA at the beginning of the 1990s, along the terms of the above trade-off:

**Figure 11: Purchasers: Trade-off Between Consumer-Responsiveness and Effective Risk Pooling**

In the US, new institutional purchasers (health alliances) were envisaged by the Health Security Act (the Democratic Administration’s plan of reform of the health system presented in 1993). The policy debate centered on the role and "size" which health alliances would take. The first issue was the relationship of the new agencies vis-à-vis health insurers. The plan envisaged that health alliances would initially take the role

\(^{77}\) Note that the ‘risk” adjusters that serve redistributive aims are conceptually different from those that are meant to limit risk-selection. The former can also be based on geographic and social variables, while the latter are based on individual data. However, the use of both types of risk-adjusters becomes more difficult and less likely to limit negative side effects of competition, when purchasers are smaller and when competition among them is more intense.
of purchasers ("demand sponsors"), but not replace insurers. However, the integration between insurers and providers in health plans was meant to progressively substitute the traditional ‘passive’ insurer. Second, the "size" of the population covered by each health alliance, and their mutual relationship was debated. The Health Security Act foresaw a role for health alliances extremely different from that originally assigned to them by Enthoven, the health economist who first suggested interposing purchasing agencies between patients and providers. It is interesting to briefly describe the debate in the light of the trade-off just explained. Enthoven (1991) favored a system based on competition on the purchaser side. He therefore favored the abolition of tax-exemptions in order to enhance price-sensitiveness of consumers, and was in favor of small health alliances competing with each other. By contrast, the Health Security Act, preferred to emphasize the aim of universal coverage, and intended to exploit the economies of scale that pooling risks in larger alliances could assure. As we can see, a balance between the terms of the consumer-responsiveness, effective risk-pooling trade-off must be found in the choice of the purchaser’s size. The optimal size of purchasers will depend on the territorial distribution of the population, on how accountable purchasers are to pressures from patients, and on the existing possibilities of re-insurance. In the US the defeat of the Democratic Party in the 1994 Senate elections interrupted the debate on health reforms. In terms of the above trade-off and relatively to the European experience, the American system continues to favor consumer-responsiveness of purchasers and does not really prevent risk-selection. Risk-selection is partially limited as insurers are allowed to fix premiums according to experience rating (which, however, makes the American health system perform very poorly in terms of "equity of access"). The consequences of incomplete coverage are also mitigated by direct government intervention (for example, through the public program Medicare that offers universal insurance to the old and chronically ill).

By contrast, in the case of the UK the 1991 quasi-market reforms gave the purchasing function to District Health Authorities and GP Fund holding schemes. The former were able to take full advantage of large risk pooling, while they largely remained unaccountable to the patients they served. GP fund holders, on the contrary, were much smaller in terms of population covered and could be abandoned by unsatisfied patients. However, their ability to manage risk was never really tested (Matsaganis and Glennester, 1994), because they always had the possibility to shift more expensive patients to HAs, and to game the system (for example, by recurring to emergency admissions). In 1997 the new Labour government initiated a reform plan replacing HAs and GPFHs with Primary Care Groups (PCG), later renamed Primary Care Trusts (PCT) on the purchaser side. PCT are commissioning and service provider agencies with an assigned pool of patients larger than that formerly administered by GPFHs, but smaller than HAs; they should promote a new coordination of community and primary care services. However, the key features of the 1991 internal market reforms are still preserved.

One final observation on this trade-off: when purchasers integrate downstream and start directly offering a wide range of health services (as in the case of American Health Maintenance Organizations, or of health purchasers in several countries of Latin America), the incentive for risk selection can be limited, as cost reductions may be achieved by directly intervening in the process of service provision. However, the drawback of this solution is that it may cause restrictions in the availability and the quality of treatment, as the next trade-off explains.

**Efficiency Versus the Right to Equal Treatment**

The second trade-off we discuss is between the objective of promoting more productive efficiency through competition and contracting out of services, and the objective of maintaining similar health service standards everywhere.

The RAP health reforms initiated in many countries are based on the view that one of the more important factors explaining providers’ inefficiency and scarce consumer-responsiveness is the lack of competition. Thus, RAP reforms have introduced new financial incentives and they have encouraged competition.
However, as competitive pressure grows, in the longer term we are likely to observe increasing horizontal and vertical segmentation in the health market. All service markets are characterized by the fact that product quality is very important and preferences are heterogeneous: therefore, providers will try to relax the competitive pressure, by differentiating their services from those of their potential competitors. Horizontal (variety) differentiation, such as geographical differentiation, is exactly what reformers seem to intend: by horizontally differentiating their services, providers specialize in those services where they enjoy a comparative advantage in terms of knowledge of market conditions or technology. As a result, patients are given wider choices, and efficiency is enhanced. Intense horizontal segmentation also brings the less desirable result that price competition is reduced (as each provider concentrates on a particular market niche), and therefore that the endogenous pressure for cost containment could be quite modest in the long term. Incentives for cost containment, however, may still be introduced by appropriately designed payment systems and contracts between purchasers and providers, as we previously discussed.

By contrast, vertical (quality) segmentation is not at all the result that reformers of public services would like to see. By vertically differentiating their products, providers attack market sections characterized by different ability to pay, focusing either on low quality, low cost services, or on high quality, high cost services. In the presence of both private competitors and a public provider bound to universal coverage, vertical segmentation can degenerate in cream skimming, whereby private providers attach only the profitable part (‘the cream’) of the market, leaving the unprofitable part (‘skimmed milk’) to the residual public provider. In the health sector, the “cream” could be constituted by elective services, or in terms of individuals, those able to pay, and the “skimmed milk” would be constituted by emergency or chronic care, or it is identified with the old, chronically ill, or simply with the poorer, less educated segments of the population.

It is claimed that cream skimming leads to a degradation of public services, for two reasons:

- By distilling out the profitable sections of the market, private providers may interrupt the cross-subsidy that sustained higher standards for the unprofitable sections;
- The ‘cream’ is usually also the informed part of the market that can put pressure for a general service improvement (in Hirschman's terminology, by exerting ‘voice’, rather than ‘exit’). Once they are gone, the public provider is left with a pool of weaker and less informed clients, which is willing to accept worst service conditions.

Phenomena of cream skimming inevitably occur when the health market is liberalized. Regulation of contracts between purchasers and providers can correct risk-selection, undertreatment and dumping of patients, and vertical segmentation, and re-establish a balance. Unfortunately, such corrections entail tempering individual incentives and restoring central control.

78 In health, the dependency of demand upon supply may reinforce this result in health.
Another important trade-off is between using payment systems to make providers more cost-conscious, and at the same time to improve their quality and avoid ‘dumping’ of more costly patients.

In order to create the proper incentives for cost-control, both purchasers and providers are increasingly subject to prospective funding, such as capitation payment system. Reformers would also like providers to keep high standards and not to make profits by selecting less costly patients, or by denying necessary care. Unfortunately, there is no single reimbursement system that is able to provide full incentives for cost containment, without creating the incentives for quality skimming and for risk selection, and the introduction of prospective payments for providers undoubtedly increases their incentive to under-treat (Chalkley and Malcomson, 1998). We can graphically represent the position where the different providers’ reimbursement systems stand along the lines of this trade-off:

![Figure 12: Providers and Type of Reimbursement](image)

As the above figure shows, moving from fee-for-service towards DRG-based reimbursement, cost-and volume contracts, capitation funding and block contracts progressively increase the part of providers’ payment that is fixed ex-ante. The incentive to contain costs increases, but so does the incentive to skim on quality and risk select by dumping potentially costly patients.

What is the significance of the trade-off just explained for equity? The link is not immediate. In order to ascertain the impact on the poor, one has first to determine how payment systems have changed for providers at different levels of care and for different services, and then how these

---

79 If demand is well informed, responds to quality, and there is excess capacity of provision, then the optimal quality of treatment can be reached also with prospective payments (Ma, 1994, Chalkley and Malcomson, 1998). However, these conditions are unlikely to be realized in the health sector.

80 Moving towards more prospective payments increases the degree of ‘supply-side cost sharing’ and therefore encourages providers to reduce cost and limit intensity of treatment. A purely prospective payment would consist of a block-contract, where the budget received by the provider is determined ex-ante and completely independent of the number and the cost of the cases treated.
changes would affect availability and quality of services for the poor, given their health seeking behavior. For instance, consider the case where the payment system for primary care doctors is changed from salary and activity-based towards capitation, and for hospitals from input to activity and cost based (as it happened in the 1990s in Poland and several other former socialist countries). According to the above analysis, in such circumstances, one is likely to observe an increase in the relative share of hospital expenditure and a reduction in that of primary care. A relative decrease in primary care vis-à-vis hospital care may well produce a negative impact over the poor, particularly those dwelling in rural areas, because they must sustain higher costs to reach hospitals in the cities, and rely more on primary care (see Tables on distribution of benefits presented in section 2; in particular, evidence from India, Figure 3).

CONCLUSION

We have shown that RAP reforms are composed of a vector of elements, and we have explained what those elements are. We have also highlighted that the different RAP reform components are likely to impact upon equity in a different way. The new resource allocation and prioritization criteria are likely to create a potentially positive equity impact, by making the distribution of resources across different parts of any country more transparent, need-based, and by reorienting public expenditure on services that benefit the poor. By contrast, for the purchasing and payment system components of RAP reforms, we have identified a few positive experiences, but we have also underlined that results on their equity impact are so far inconclusive and mixed. One key concern is that these new RAP mechanisms, introducing competition among purchasers, contracting with the private sector, and output or outcome-related payment systems, would raise the possibility of introducing severe distortions in providers’ behavior to the detriment of the poor. Specifically, purchasing, contracting and payment system reforms meant to reduce costs, increase efficiency-or improve quality of services may in fact lead to widen quality differentials across services, and increase providers’ incentive to cream skim, risk select and under-treat poorer or more vulnerable patients. However, in the absence of any solid empirical evidence, the arguments in favor or against remain primarily theoretical (see discussion of the tradeoffs).

The new allocations mechanisms and incentives introduced by RAP reforms, characterized by more transparent geographical distribution of resources, output/outcome orientation, arms-length relationship with providers, etc. has the potential to really improve the current “way to do business” in the health sector. Probably, their full potential is not yet exploited. However, implementing RAP reforms entails intensive resource as well as human capital investments in order to be appropriately implemented and monitored, and to prevent potential adverse effects. It also requires governments able to find within the RAP reforms’ tool-kit a mix of interventions and mechanisms able to maximize the reforms’ positive impact, and to control adverse effects, particularly on the equity dimension. The ability to implement these investments and this

81 If drug expenditure is reimbursed to patients for inpatient but not for outpatient care, such effect will be more pronounced.

82 Consider for instance the use of capitation-based funding for health purchasers and/or providers. This is an extremely powerful tool to contain cost and to improve transparency in the distribution of resources across providers, but it is also likely to induce them to risk-select patient and to under-treat them (see Mills et al. 2000).
steering/oversight in turn presupposes a health financing framework and other preconditions which may not be yet feasible in the poorest, capacity constrained countries.

We also repeatedly underlined that it is necessary that ongoing and future reforms be subject to a more rigorous monitoring and evaluation process. It is striking to find out that quite radical health systems reforms, not only in Africa or South Asia, but also in Central Europe or even in the UK, over the 1990s were in fact implemented without any systematic monitoring system which could allow an ex-post rigorous evaluation of their impact. We recommend that, wherever it is possible, future RAP and other health systems reforms be phased in progressively and systematically collecting baseline and post-reform data. More attention should also be devoted to collecting data disaggregated by gender, age and socio-economic group.

In conclusion, RAP reforms were not originally justified because they intended to improve services specifically for the poor. However, the above discussion suggests that any health reform that succeeds in improving quality and effectiveness of health services in general, and particularly public services, without reducing accessibility, it is also likely to have ipso facto a strong positive impact on the poor.

However, by adjusting the capitation formula in order to cross-subsidize care for the more vulnerable/disadvantaged groups, such adverse effects can be controlled or at least limited.
ANNEX 1 UNDERSTANDING THE LINK BETWEEN SOCIO-ECONOMIC DEPRIVATION AND HEALTH EVIDENCE FROM INDUSTRIALIZED COUNTRIES

Both theoretical understanding and empirical evidence suggest that poverty and poor health are strongly associated, and that health improvements and poverty reduction influence each other through several channels: physiological, behavioral and psychological. This Annex summarizes the different channels through which poverty has an adverse impact on health, focusing mostly on literature from industrialized countries. Recent literature has also started investigating the reverse causal relationship, from poor health to poverty (for a review of this literature, see Belli and Subramanian, 2001). This literature is showing that improving health indicators, especially among the poor, has in the medium-long term a strong positive effect on economic growth. In other words, according to this view, the achievement of improved health outcomes among the poor (together with improved education indicators) tells us that some fundamental improvement is occurring within a community or country, and that improvement is most likely to herald more robust economic growth in the future.

First noteworthy conclusion from the literature on the socio-economic determinants of health is that income and income distribution plays a role at least as important as access to health services in determining inequalities in health status across individuals.

In the literature on industrialized countries, there are two views within which the connection between income and health is captured. The first, more recently investigated, is the "relative-income" view. Kawachi et. al (1999) identify three pathways through which income distribution affects individual health. First is through reduced access to life opportunities, material resources and opportunity structures (see Kaplan et. al. 1996). For instance, in more unequal societies the ‘pooling’ of resources, that could finance public services such as health care systems and education, is difficult to achieve, since the richer segments tend to typically use privately funded and provided services. Cross subsidies across socio-economic groups become difficult to politically sustain, thus affecting the accessibility and quality of public services. The second set of processes is linked to decreasing social capital that manifests through factors such as erosion of social cohesion, increased social exclusion and conflict (see Putnam, 1993, 2000; Kawachi and Kennedy, 1997; Wallace and Wallace, 1997; Wilkinson, 1996). Thirdly, more direct psychosocial pathways such as hopelessness, lack of control, and loss of respect arising as a consequence of inequality have been identified as impacting upon individual health (see Wilkinson et. al. 1998).

The second is the "absolute income" view. Here, the theoretical understanding suggests that health improvements and poverty reduction influence each other through several channels: physiological, behavioral and psychological.

**Physiological**

Being poor limits the availability and the ability to take advantage of the basic pre-requisites of health, such as food, water, clean and warm clothing, safe shelter, education and access to sanitation and health care services. Inadequate food, lack of access to water and sanitation, poor
housing and homelessness, poor working environment, lack of education, more difficult access to health services, are in fact all daily dimensions of poverty. Accessibility of health services is only but one of the several factors that may influence the poor 's health status and the observed health inequalities.

**Behavioral**

Poverty may result in ill health through behavioral consequences in two ways: 1) adoption of high-risk behaviors; and 2) restriction of individuals' subjective choices. High-risk behaviors, such as smoking, are more prevalent among poor than among rich. Similar differences in alcohol intake and drug addictions have been found among poor and rich. However, evidence from developed countries suggests that the behavioral risk factors may not explain the entire effects of socio-economic status (SES) on health. First, longitudinal studies have shown that the larger part of the social class gradient in heart disease is unexplained by behavioral risk factors (Rose, 1995) and that SES remains a statistically significant independent risk factor even after adjusting for the behavioral risk factors (Davey Smith et al, 1996 and 1998). Second, there are class gradients in diseases for which we have no clear knowledge of behavioral risk factors. The diseases unrelated to smoking, for example, have as steep a social gradient as the ones that are related to smoking. Poverty also permanently reduces the ability to adopt healthy life-style, related to the lack of information/education and to the persistence of wrong life-styles acquired during young age.

**Psychological**

The poor suffer from both the direct effects of material deprivation as well as the effects of psychological, social and emotional deprivation within their societies. The material insecurity itself is a source of stress. Evidence from factory closure studies shows that unemployment leads to deterioration in health and that much of the deterioration in health started even before people became unemployed, soon after redundancies were announced (Iversen and Klausen 1981; Mattiasson et al, 1990; Ferrie et al, 1995). Moreover, other studies show a striking connection between the extent of social deprivation and the crime rate. At least part of the deterioration in health and the greater propensity to be violent may be related to psychosocial stress/aggressivity related to job insecurity, and increased deprivation.

**Health inequalities and their link with health systems’ characteristics**

Finally, we mention a recent study on the causes of health inequalities and their link with socio-economic deprivation and health systems’ characteristics in industrialized countries. The study is by Van Doorslaer et al. (1997). They present evidence from nine OECD countries, and in their first exercise they regress self-assessed health status (cardinalised according to a standard lognormal distribution), on income. Their result is that the USA appears to have the highest level of health inequality, followed by the UK, the Netherlands and then all the other countries. Inequality is statistically significant in all countries. In their second exercise, Van Doorslaer et al. regress health inequalities on:

1. Total pro-capita health expenditure
2. Public share in total health expend.
3. Total GDP per capita
They find that only the last term, income inequality, seems to have a positive statistically significant association with health inequality, whilst income per capita, total health expenditure and even the public share seem not to matter. His finding seems to confirm the importance of the “relative” pathway through which income and income distribution may affect poor health.
As we discussed in the text (section 2), for an equity assessment concerning the distribution of any resource or asset in society it is critical to first characterize the “benchmark distribution”. In general, the benchmark distribution is not perfectly egalitarian, and it can be different according to each of us. However, for everything that is considered as a human right, matters are different and more straightforward. By definition, individual human rights, unlike assets and resources, are inviolable, inalienable (one cannot buy or sell his/her right to vote, or his/her right to free speech) and must be equally enjoyed by all. Thus, perhaps the first question to ask is the following: “Shall we consider health as a human right?”

I believe that the critical distinction here is between those that consider as proper human rights only the so-called Negative individual Rights, and those that extend the realm of rights to include also Positive individual Rights. The former claim that special protection should only be guaranteed to a few core rights, such as the right to safety, to freedom of expression, to non-discrimination, etc. Those are the original human rights that were first articulated in the period of the war for Independence in the United States and of the French Revolution. According to this view, governments should protect the above limited group of Negative Human Rights, and for the rest not interfere with individuals’ decisions. More recently, comparable arguments have been forcefully re-proposed by the noted American philosopher Nozick (1974), advocate of the so-called Minimal State Theory. According to Nozick, what matters to judge over the "fairness" of a society and over the allocation of resources resulting from social interaction, is whether or not individuals’ Negative Rights have been respected, and whether the "rules of the game" are the same for everybody regardless of the initial conditions. Equity is an attribute of the process, and not of the end state.

On the other hand of the political spectrum are those who claim that it is what each person can actually be or do in a specific society that matters, and not so much what the government formally allows each of us to be or do. Against a formal view of freedom typical of the Negative Rights ‘school’, which perceives freedom mainly as the absence of interference from others, the focus here is primarily on end-states and Positive Rights. These include such rights as that to adequate nutrition, to housing, to education, to health services, to political representation, to the possibility of employment, etc. The claim is that a society that respects all Negative Rights may in fact be extremely unjust, if it de facto excludes part of its citizens from the possibility of flourishing as

---

83 Such rights were considered as "natural", bestowed on each individual by God, first in the 17th century by Iusnaturalist philosophers, and then in the 18th century by the Founding Fathers of the U.S.A. and by French Enlightenment philosophers. In those times, the government mainly meant “the king”, and so those theories were actually first meant to protect each person from the arbitrariness of royal power.

84 It is interesting to see how some of the key dividing lines that would subsequently characterize the political spectrum up to the present days, described in the text, were already apparent in the interpretation given to the ideals of Liberte' and Egalite' by the different political factions that fought for power in France in the period 1789-1795.

85 In the period of the industrial revolution, among all the Negative Rights a prominent place was assigned to property rights and to the right of free economic initiative.
individuals because of material constraints. The typical 19th century Marxist example is that of a society that protects Negative Rights, including private property and economic freedom to the fullest, but where most people have no choice, but to sell their labor under appalling conditions in order to survive. According to this second viewpoint, the answer to the question: “Equality of what?” is: “Equality in Positive Rights and, to an extent, in end-states”.

In development assessments, the shift towards putting people’s Positive Rights at the center came to be recognized with the advent of the ‘Basic Needs’ approach (ILO, 1976; Streeten and Burki, 1978, Streeten et al., 1981). This approach emphasizes the priority of meeting the basic needs of people, and the content of these needs is seen to include nutrition, health, shelter, water, sanitation, education and other essentials (Desai, 1990). More recently, building and elaborating over such theories and over Rawls’s “Theory of Justice” (1971), the noted Indian economist and philosopher Amartya Sen argued that development should be conceptualized and assessed as peoples’ ability to do things that they have a reason to value, what he called the ‘capability’ approach (Sen, 1973, 1984, 1985, 1992). Critical to the capability argument is the notion of human freedom (Sen, 1992, 1999), and what characterizes the capability-position is again a peculiar view of freedom. Sen, distancing himself from Marx, endorses the ‘intrinsic’ as opposed to the ‘instrumental’ view of freedom (Sen, 1989). The intrinsic view values freedom for its own sake, whereas the instrumental view considers freedom to be important merely because of its significance for other achievements. However, distancing himself also from the Negative Rights tradition, Sen specifically proposes that assessments be based on the ‘positive-intrinsic’ type of freedom. Thus, Sen’s reply to the question: “Equality of what?” is: “Equality of capabilities across individuals in society” (Sen, 1993). For equity-based social assessments dealing with developing countries, Sen proposes a freedom-based examination of the level and distribution of relatively elementary, yet important, capabilities, of which health is a key component.

The question: “Equality of what?” and the distinction between the Negative and the Positive Rights ‘schools’ of thought is so important because the broader is the set of Positive Rights that society is prepared to protect, the more the government must encroach upon individual’s Negative Rights in order to do so. In other words, moving towards more equality over some dimensions, in general requires moving away from it over some other dimensions. For example, increasing the range of social entitlements to fulfill Positive Rights may entail restricting

86 The first to open this perspective on justice was Jean-Jacques Rousseau (1712-1788), in his work: “The Social Contract” (1762). Rousseau imagined that the rules governing a just society must be determined by referring to a State of Nature, where individuals still do not know which position they will occupy in society, a thesis re-affirmed in modern times by J. Rawls (1971). The first political leaders who referred to Rousseau's ideas were the Jacobin leaders, during the revolutionary government in France (1792-1794).


88 Sen (1987) argues that a person’s well being is characterized not only by the state (functioning) in which a person lives or is living, but also by the existence of a range of alternative states from which the person can choose (capability). In empirical studies, it turns out that it is in fact extremely difficult to distinguish between capabilities and achieved functionings.

89 Sen’s thesis has also been widely used by the UNDP for its ‘human development’ agenda. See the collection of papers in UNDP (1997). Also see Anand and Sen (1994a, 1994b, 1995).
individual economic freedom, or at least some individuals' faculty to spend their money in the way that they see fit. Or, imposing to individuals a level of utilization of health services compatible with achieving the right to health would entail violating consumer sovereignty\textsuperscript{90}. Vice-versa, respecting individual preferences would entail violating equality in utilization of health services, and even more equality in health, because, given that demand income elasticity is greater than one for health services, without government intervention we would expect higher utilization and also a higher share of income spent on health services among the rich than the poor. In synthesis, there are trade-offs between different values and we, as members of society, are called to choose among them. For human rights, we should aim for equality of outcomes, whereas for other dimensions inequalities and, given the pervasiveness of trade-offs, indeed ‘trade’ between conflicting aims is to be accepted.

In the literature more specific to health, Culyer and Wagstaff (1993), following Sen and a tradition deeply rooted in Europe, claim that health ought to be considered as a Positive Individual Right, and thus they conclude that equity demands that all equally enjoy it. The same position is echoed in recent official policy documents\textsuperscript{91}, and, for subsets such as child health, in international treaties\textsuperscript{92}. Alternatively, some consider that the realm of individual Positive Rights should be limited to a minimum level of entitlements against the great risks of life, among which is illness\textsuperscript{93}. The emphasis here in on services, rather than outcomes, and on the limits of the human rights’ perspective in health to a circumscribed set of services. The extent of protection beyond the minimum should instead vary according to the level of socio-economic development, and according to the social preferences that characterize each society.

Is it possible to reconcile the two views? Culyer and Wagstaff (1993), argue that health is a Positive Right because it is one of the more important material preconditions necessary for each of us to flourish as an individual. In this respect their argument is convincing. Unlike other rights, however, being healthy or reaching old age in good health is not something that governments can equally guarantee to all. Thus, a first obvious qualification of the right to health, as Gakidou, Murray and Frenk (2000) point out, is that the maximum achievable target may in fact be to equate individual health expectancies (for instance, across socio-economic groups). Variations in healthy life spans that are simply due to chance cannot be avoided. More importantly, recent research shows that the society’s socio-economic structure (its level of income inequality, for instance), and other causes external to the health sector are likely to account for the

\textsuperscript{90} Given the influence of all the other socio-economic determinants, achieving equal health expectancies across socio-economic groups would actually entail forcing a disproportionate use of services by the poorer members of society (see Annex 1).

\textsuperscript{91} See, for example, WHO World Health Report 2000.

\textsuperscript{92} For instance, the Convention on the Rights of the Child, ratified by every country with the exception of the United States and Somalia, recognizes health among the fundamental rights of all children (Art. 24).

\textsuperscript{93} The emphasis here is on financial protection. This implies that public support should be stronger in those situations, such as in absence of family support, when people are generally unable to cope, if left alone. The British Social Security System, and in particular, the pension system, was articulated along these lines. It had from the beginning the main function of providing only a minimum safety net.
distribution of health status within a society, probably as much as, or more than the quality of medical care per se. This finding has two implications. First, that the right to health cannot be guaranteed unless other rights, such as the right to adequate nutrition, to healthy housing and healthy working conditions are also guaranteed. Second, one should consider that perfect equality of health expectancies most probably couldn’t be achieved, unless all people enjoy the same socio-economic condition and the same life-style.

Thus, as we pointed out above discussing Negative vis-à-vis Positive Rights, it is a matter of trading-off the aspiration to equal health with other aspirations. But the definition of a human right excludes the possibility of trade-off. Thus, here we find a second necessary limitation to the idea of health as a human right. For health (as it is for housing or education), there is a continuum between the state of perfect health and that of death, and each society needs to draw a line between what is reasonable to consider as an individual right and what ought to be considered differently. Following this perspective, we may for instance argue that maternal and child health are indeed to be considered as human rights, so that equity demands that safe-motherhood and equality in health outcomes among children be pursued to the fullest (by implementing strongly pro-poor interventions), whilst certain differences in adult health and in utilization of health services across adults should be accepted, because eliminating them would severely limit other rights, that we consider more important.

Finally, we can further qualify the human rights approach to health by taking an empirically oriented perspective (see, for example, Willis, 1993, Waters, 2000). Regardless of the conclusions from a philosophical point of view as to what extent a human right to equal health should be protected, what governments can actually measure and somewhat control are access to and utilization of health services, rather than health status per se. This constraint can be rephrased by referring to the distinction between inputs, outputs and outcomes, which is so important in health. Equity objectives might be posed in terms of outputs or outcomes. However, the public sector may have control only over some inputs, but not fully on outputs and even less on outcomes. Thus, to make the right to health “operational”, ‘all’ governments can do is to enhance access and utilization of health services by the poor, promote healthy behaviors and impose universal respect of certain health protection measures in production and consumption (at home, in the workplace, on the roads, etc.). Then, our equity focus or assessment should be more on the utilization of health services or health care, rather than on health outcomes. That is the focus we have decided to take in the text.

94 A related issue is to determine along which dimension it is best to redistribute inputs (resources) and outputs. Is it better to sharpen targeting of health services, or to achieve better education for the poor, or to build safer roads in the rural areas, or greater equalization of income and wealth through taxation? Interventions channeled through sectors different from the health system may in fact achieve much more significant results in reducing health inequalities than those provided directly within the health system (see, for example, Hammer, 1998).

95 Of course, the extent of health and health care that constitutes an individual positive right, or of “the minimum amount of health care entitlements” is not fixed, but evolves over time. We can draw a parallel with education: the extent of the right to education has progressively been extended over time, from simple elementary school education, to the present compulsory school curriculum. Yet, part of the higher education is still allocated “unequally” in society, according to preferences and merit.
Beginning in the late ‘60s, early ‘70s, a concern for the provision of primary care services, and against an exclusive focus on hospitals and specialized care, progressively emerged in developing countries, as exemplified in the Alma Ata declaration (WHO, 1978)\(^{96}\). Then, during the ‘80s and ‘90s, a consensus progressively emerged within the health research community to utilize burden of disease and cost-effectiveness as the principal criteria \(^{97}\) to determine priorities for public health intervention. On the basis of such criteria, the World Development Report (1993) advocated that priority for public funding be assigned to a minimum ‘package’ of cost-effective services (Basic Benefit Package, or BBP).

The 1993 WDR’s recommendations gave rise to several thorny issues concerning the definition of priorities for public financing and provision, which have not yet been resolved. In the debate that followed the publication of the WDR, two among several others points of critique centered around the WDR’s lack of attention for:

- **Implementation/incentives issues.** Theoretical cost-effectiveness (MICE) computations used in the WDR were abstract, derived in ideal settings, and ignored a host of complicated feasibility/implementation/incentive issues. Filmer et al. (1998) showed that if one considered the evidence on actual public sector cost-effectiveness (PSCE) from developing countries, one would reach estimates of cost effectiveness several orders of magnitude higher (worse), than those estimated by theoretical MICE computations. Thus, part of the international research effort after the publication of the WDR has tried to articulate and test new ways to improve efficiency and quality of public delivery in developing countries.

- **The role of the private sector, and the public-private interaction.** Policy recommendations in the WDR 1993 failed to recognize the simple reality that health interventions by the public sector do not take place in a vacuum. The recognition of the private/public differential ability in providing/financing different health services has lead to revisit the criteria for determining priorities. As Hammer (1997, p. 48) writes: "Government investment, like any other government intervention, should be justified in terms of the social benefit the project will have over and above that which would occur without public sector involvement. For any investment opportunity, the focus of analysis should be on the difference between social and private benefits- not on the costs and expected returns to private goods themselves". Whenever the public sector offers or

---

\(^{96}\) Traditionally, public intervention in health was focused on expanding the delivery system, through physical infrastructure investment and increase in the number of doctors and nurses.

\(^{97}\) Medical intervention cost-effectiveness (MICE), the principal criterion, ranks interventions according to their ability to achieve a certain outcome (lives saved or disability-adjusted life years gained) per unit cost, according to best practice protocols. Burden of Disease (Body) studies provide estimates of how important different health disorders are (for instance in terms of Daly’s lost) to the total BD Criteria for essential package was, in the end, a combination of magnitude of burden + MICE of interventions. However, Gwatkin’s work on disease burden in poorest groups showed that these two criteria led to different conclusions.
subsidizes health services that the private sector is already offering, this is most likely to influence the way the latter operates\(^98\).

This angle of looking at priorities for public intervention, sometimes called the welfare economics perspective, bears important consequences for equity. For instance, even if most primary care and outpatient services are in general more cost-effective than most inpatient care services\(^99\), the potential for cost-recovery is much greater for the former than the latter. Catastrophic episodes that require inpatient care are in general the less predictable health events and the more costly to treat. In the absence of private insurance coverage and/or direct subsidization of hospital care, the majority of the population could not afford to treat them\(^100\).

A recent contribution by Musgrove (1999) tried to reconcile the principles of welfare economics and those of cost-effectiveness. Musgrove provides a decision tree on the rational use of public financing in the health sector based on the two set of principles\(^101\). His contribution revamped the debate on prioritization criteria, and other scholars have forcefully proposed again the argument that cost-effectiveness should not be a criterion used to determine public resource allocation priorities\(^102\). Recently, Filmer and Hammer (2002), using the criterion of welfare loss/market failures, plus equity considerations, recommend that priority should be given to (i) pure public goods, (ii) services with large externalities (such as water & sanitation, immunization services), and (iii) services for catastrophic illness.

---

\(^{98}\) There may be a "crowding-out" or "crowding in" effects, so that only changes in net consumption are relevant, as well as variations in the quality of services offered by the private sector. Moreover, by influencing demand for curative services through public health, information and preventive campaigns, the public sector influences the size of the market and thus market equilibria.

\(^{99}\) Medical treatments for most catastrophic illnesses tend to occupy low levels in priority scales based on MICE analysis, as they are relatively more expensive and rarely resolving (see Hammer and Berman, 1995).

\(^{100}\) In the language of welfare economics, the private market equilibrium is characterized by market failures that entail welfare losses.

\(^{101}\) Mangrove's (1999) decision tree starts with the welfare economics questions, by asking if the service proposed for public funding is a public good, or it bears significant positive externalities. If it does, the next step is to rank it vis-à-vis other services with the same characteristics in terms of cost effectiveness or benefit-cost analysis. In contrast, if the service is question is a private good, or does not entail significant externalities (such as most curative care services), the decision tree asks whether risk of catastrophic costs are involved and whether the proposed beneficiaries are poor or not. Again, all interventions for which the answer is positive are ranked according to cost-effectiveness.

\(^{102}\) Jack (2000) points out that cost-effectiveness is a criterion to compare alternative services, once it has already been decided that those services are worth of social protection and therefore that the government should subsidize them (because they are non-rival or non-excludible, because there are significant externalities involved, or to correct insurance market failures, or because the services create a pro poor impact, or, more in general, because health is considered as a human right, or a merit good. Merit goods are goods that the government should subsidize beyond the level that individual preferences would dictate. The concept was initially developed by the German school of Public Finance, and later introduced into the mainstream of Public Economics by R. Musgrave.
REFERENCES


Belli, P., 2001b, Ten years of health reforms in the ECA region. Lessons learned and options for the future, WP, Center for Population and Development Studies, V.11 n.6, Cambridge. MA.


Berman, P. and T. Bossert, 2000, A Decade of health sector reform in developing countries: what have we learned? Paper prepared for the DDM Symposium, Appraising a decade of health sector reforms in developing countries, March 15th Washington DC.


Bossert, T., 2000a, Analyzing the decentralization of health systems in developing countries: Decision space, innovations and performance, *Social Science and Medicine*, 47: 1513-27


Ensor, T. and J. Langenbrunner, 2000, Allocating resources and paying providers, World Bank (Draft).


Evans, R., 1974, Supplier-Induced Demand: Some Empirical Evidence and Implications, in M. Perlman (eds.), The Economics of Health and Medical Care, London. McMillan.


Fox D.M., and H. Leichter, 1993, The Ups and Down of Oregon’s Rationing Plan, Health Affairs, summer 1993,


Gilson L, van den Heever A, Brijlal V, 1997. DHS Development in the North West Province : District Financing and Management , center for Health Policy , Department of Community Health , University of Witwatersrand


Gertler, P., and J. Molyneaux. 1990. Pricing public health services: Lessons from a social experiment in Indonesia: RAND, Santa Monica, California


Gupta, S. Verhoeven, M., Tiongson, E., 2001, Public Spending on Health Care and the Poor, IMF Working Paper /01/127


Hanson K., Ranson, K., Oliviera-Cruz, Mills, A., Constraints to scaling up health interventions: A conceptual framework and empirical analysis, CMH WP, no WG5: 14.


Hanson, K., Berman, P., Assessing the Private Sector: Using non-government resources to strengthen public health goals, WP, Data for Decision Making Project, IHSG, Harvard School of Public Health.


Jack, W., 2000, Public spending on health care: how are the different criteria related? A second opinion, Health Policy.


Lake, S., Mtonga, S., Nakamba, P., 2002, Strengthening needs-based resource allocation


Narayan, D. et al., 1999, Global Synthesis: Consultation with the Poor, World Bank: Washington DC.


Palu, T., 1999, Hospital Sector in Transition Baltics Case Study. (Draft.) World Bank, Washington, D.C.


Sánchez F., Núñez, J., 1999, La Ley 100/93 Sí Llegó a las Personas de Menores Ingresos”, La Revista VíaSalud, third trimester, Santafé de Bogotá, Colombia.


Sen, A., 1985, Commodities and capabilities, Amsterdam: North Holland


81

——1998b. Reducing Poverty in India: Options for more effective public service

——1999c. Under-the-Table Payment and Health Care Reforms in Hungary. Observatory Case Study (Draft). Health Services Management Training Centre, Semmelweis University of Medicine, Budapest, Hungary.


——2000b. World Development Indicators.

——2001a. HNP Chapter for PRSP Sourcebook.

——2001b World Development Indicators.

——2002a, Georgia Public Expenditure Review, The Health Sector.


World Health Organization, Geneva:


——1999b, Poverty and Health: Report by the Director General, document EB105/5 presented to the 105th session of the WHO Executive Board.


——2001a, Health for All Database.


The Economics of Priority Setting for Health Care: A Literature Review

Katharina Hauck, Peter C. Smith and Maria Goddard

September 2004