ANALYZING THE DECENTRALIZATION OF HEALTH SYSTEMS IN DEVELOPING COUNTRIES: DECISION SPACE, INNOVATION AND PERFORMANCE

THOMAS BOSSERT
Harvard School of Public Health, 665 Huntington Ave. I-1210, Boston, MA 02115, U.S.A.

Abstract—Decentralization has long been advocated as a desirable process for improving health systems. Nevertheless, we still lack a sufficient analytical framework for systematically studying how decentralization can achieve this objective. We do not have adequate means of analyzing the three key elements of decentralization: (1) the amount of choice that is transferred from central institutions to institutions at the periphery of health systems, (2) what choices local officials make with their increased discretion and (3) what effect these choices have on the performance of the health system. This article proposes a framework of analysis that can be used to design and evaluate the decentralization of health systems. It starts from the assumption that decentralization is not an end in itself but rather should be designed and evaluated for its ability to achieve broader objectives of health reform: equity, efficiency, quality and financial soundness. Using a “principal agent” approach as the basic framework, but incorporating insights from public administration, local public choice and social capital approaches, the article presents a decision space approach which defines decentralization in terms of the set of functions and degrees of choice that formally are transferred to local officials. The approach also evaluates the incentives that central government can offer to local decision-makers to encourage them to achieve health objectives. It evaluates the local government characteristics that also influence decision-making and implementation at the local level. Then it determines whether local officials innovate by making choices that are different from those directed by central authorities. Finally, it evaluates whether the local choices have improved the performance of the local health system in achieving the broader health objectives. Examples from Colombia are used to illustrate the approach. The framework will be used to analyze the experience of decentralization in a series of empirical studies in Latin America. The results of these studies should suggest policy recommendations for adjusting decision space and incentives so that localities make decisions that achieve the objectives of health reform.

INTRODUCTION
Decentralization has been promoted by advocates of health sector reform in developing countries for decades. Viewed initially as an administrative reform which would improve efficiency and quality of services and later as a means of promoting democracy and accountability to the local population, decentralization was seen by many advocates as a major reform in and of itself. Despite this advocacy, until very recently only a few nations have actually adopted and implemented decentralization reforms. This lack of experience is reflected by the few empirical studies which examine the actual impact of decentralization*. There has been no systematic study using a common analytical framework to examine the relationship between processes and types of decentralization and actual outcomes or performance in the health sector.

OBJECTIVES OF THIS ARTICLE
The objective of this article is to develop a comparative framework to analyze the effectiveness of decentralization for reaching the goals of health systems in developing countries. This framework will be used by the author in a series of forthcoming empirical studies in selected developing countries that have sufficient experience with decentralization to evaluate performance.

A comparative analytical framework should provide a consistent means of defining and measuring decentralization in different national systems. It should help define the different degrees of decentralization and the mechanisms that are used to influence and control decisions at local levels. Secondly, the analytical framework should clarify how decen-

*Reviews of this empirical literature are: Peterson (1994), Prud’homme (1995), Bossert (1996), Collins (1996) and Cohen and Peterson (1996). Examples of this literature include studies of decentralization in Papua New Guinea (Kolehmainen-Aitken, 1992; Campos-Outcalt et al., 1995), Mexico (Gonzalez-Block et al., 1989), Brazil (Tendler and Freedheim, 1994), Colombia (World Bank, 1994), Chile (Bossert, 1993), Bolivia (Holley, 1995) and the United States (Altman and Morgan, 1983).
Centralized systems differ from centralized ones, in both process and outcome terms. What different choices will result from local decision-making compared to centralized decision-making? Thirdly, the framework should develop performance indicators to evaluate the impact of different choices made by local decision-makers. These choices should be evaluated by assessing how they contribute to the general goals of health system reform: improving equity (including universal coverage, access and solidarity), efficiency, quality and financial soundness*. In this way, we can view decentralization as a means toward the ends of broad health reform, rather than an end in itself.

In this paper I will first review the four major analytical frameworks that have been used by authors who address problems of decentralization in the health sector: (1) public administration, (2) local fiscal choice, (3) social capital and (4) principal agent approaches. I will discuss the strengths and weaknesses of these approaches. Then, I will propose using the principal agent approach as a general framework for analysis. I will develop this framework by introducing the concepts of “decision-space”, “innovation” and “directed change”. I will illustrate some key issues with examples from Colombia, one of the limited number of countries with several years experience of implementation of decentralization.

In brief, the proposed approach develops a comparative definition of decentralization which focuses on the range of choice that is available to local decision-makers along a series of key functional dimensions. This definition is called “decision-space” and allows us to specify and then evaluate the impact of restricting or opening the degree of local choice on financing, service organization, human resources, targeting and governance. In addition to the formal range of choice, we also need to examine the tools available to the central level to influence these choices: positive incentives and sanctions, such as providing matching grants or withholding funding. The characteristics of the local governments — such as the pool of local skilled personnel — that might influence their capacity to make effective choices should also be evaluated.

Then, the approach asks how these local authorities use the decision space and respond to the incentives: do they innovate or simply continue doing what they had done before. In some cases, we can compare this use of decision space with the “directed change” that occurs in a centrally controlled locality. Finally, the approach develops performance indicators to be used to evaluate whether different forms of decision space have allowed localities to make better decisions than were made before or by centralized localities. Figure 1 suggests the overall approach.

**REVIEW OF FRAMEWORKS OF ANALYSIS**

The following section reviews the major frameworks for analysis used in the current literature on decentralization. Our immediate objective in this review of frameworks is to determine which approach is appropriate as an overall framework for evaluating how decentralization contributes to the achievement of general health sector goals.

**Public administration approach**

The public administration approach was first introduced by Dennis Rondinelli and G. Shabbir Cheema for evaluating broad processes of decentralization in developing countries (Rondinelli and Cheema, 1983). This approach was applied to the decentralization of health systems in a seminal World Health Organization publication on the issue (Mills et al., 1990).

---

*For a discussion of health goals, see Berman (1995). For the purposes of this article we assume that these goals are the goals of the central government. In empirical studies we will evaluate the actual commitment of the central government to these internationally promoted goals.
The public administration approach focuses on the distribution of authority and responsibility for health services within a national political and administrative structure. This approach has developed a now well-known four-fold typology of different forms of decentralization: (1) deconcentration, (2) delegation, (3) devolution and (4) privatization. Deconcentration is defined as shifting power from the central offices to peripheral offices of the same administrative structure (e.g. Ministry of Health and its district offices). Delegation shifts responsibility and authority to semi-autonomous agencies (e.g. a separate regulatory commission or an accreditation commission). Devolution shifts responsibility and authority from the central offices of the Ministry of Health to separate administrative structures still within the public administration (e.g. local governments of provinces, states, municipalities). Privatization transfers operational responsibilities and in some cases ownership to private providers, usually with a contract to define what is expected in exchange for public funding.

In each of these forms of decentralization significant authority and responsibility usually remains at the center. In some cases this shift redefines the functional responsibilities so that the center retains policy making and monitoring roles and the periphery gains operational responsibility for day to day administration. In others, the relationship is redefined in terms of a contract so that the center and periphery negotiate what is expected from each party to the contract. A central issue of the public administration approach has been to define the appropriate levels for decentralizing functions, responsibility and authority (see Mills, 1994). The principal arenas are usually regions, districts and local communities.

The weaknesses of the approach are that it does not provide much guidance for analyzing the functions and tasks that are transferred from one institutional entity to another and does not identify the range of choice that is available to decision-makers at each level. There is an implicit assumption that moving from deconcentration toward privatization is likely to increase the range of choice allowed to local officials and managers; however there is no clear analysis of why this should be the case. Much of the empirical literature using this approach discusses the need to specify just what tasks or functions are assigned to each form or level, but as a framework it does not provide us with analytical tools to specify and compare tasks and functions (Gilson et al., 1994).

The strengths of this approach are that it provides a readily observable typology for identifying the institutional arrangements of decentralization. It focuses attention on the levels and organizational entities that are to receive or lose authority and responsibility.

Local fiscal choice

The local fiscal choice approach was developed by economists to analyze choices made by local governments using their own resources and intergovernmental transfers from other levels of government (Musgrave and Musgrave, 1989). It has been applied mainly in federal systems where local governments have had a history of constitutionally defined authority and significant locally generated resources. This approach assumes that local governments are competing with each other for mobile voters (who are also taxpayers) and that government officials make choices about resource mobilization, allocation and programs in an attempt to satisfy the preferences of the median voter (Chubb, 1985). Studies of federal systems have tended to find that central governments are more effective for making equitable allocation decisions (especially for assisting the poor) and that local governments more effectively utilize funds to achieve efficiency and quality objectives. One issue often stressed in this literature is the role of intergovernmental grants as substitutes for local spending, often driving out local funds for health rather than stimulating local counterpart funding (Correa and Steiner, 1994; World Bank, 1994; Kure, 1995; Wisner Duran, 1995; Carciofi et al., 1996).

There are several limitations on the applicability of the local fiscal choice approach in developing countries. First, in most developing countries, local resources are a small portion of local expenditures and intergovernmental transfers come with many administrative restrictions. It is difficult, therefore, to assume that the voter holds local authorities responsible for both the taxation, which is centralized, and the programs, which are only partially decentralized (Peterson, 1994). In Colombia, for instance, intergovernmental transfers account for over 90% of most local resources and the central government restricts local choice over these transfers. Secondly, it is difficult to assume that local authorities respond to the median voter assumptions when so many other political factors are involved in making local choices, including clientalism and patronage (Chubb, 1985). Also, voters tend not to be single-issue voters; they choose candidates for a variety of reasons, not just health care issues. Finally, the assumption of voter mobility is often unrealistic (Prud’homme, 1995).

The strength of this approach is that it focuses attention on the local decision-making and develops clear and parsimonious theoretical propositions to explain those choices. Using rational actor assumptions, it examines the incentives — both economic and political — for local decision-makers to make choices that are desired by local citizens or by central governments. The approach introduces the importance of considering locally generated revenue and the role of local politics and accountability to
the local population. While the usual assumptions of the local fiscal choice approach may not hold, the orientation toward local sources of funding and accountability to local political processes is important for generating hypotheses about how devolved systems will function.

**Social capital approach**

The social capital approach, introduced recently by Robert Putnam in his study of Italy, has generated new research in the area of decentralization. This approach focuses on explaining why decentralized governments in some localities have better institutional performance than do governments of other localities (Putnam, 1993). Putnam finds that it is the density of civic institutions — a broad range of different, largely voluntary, organizations like choral societies and soccer clubs — that create general expectations and experiences among the local population that he calls “social capital”. It is this investment in social experience that encourages people to work together rather than as autonomous self-seeking individuals and to develop expectations, reinforced by experience, that they can trust each other. He argues that it is this trust that fosters behavior that makes for better performance in local institutions.

Applied to health care, this approach suggests that those localities with long and deep histories of strongly established civic organizations will have better performing decentralized governments than localities which lack these networks of associations. In Colombia, where we do not have systematic information, anecdotal cases suggest that some regions, such as Antioquia and Valle, might have more dense social networks, which might explain why they have better performing local institutions.

The weakness of this approach is that it does not provide easy policy relevant conclusions. Areas without civic networks seem to be left out of the picture. Putnam’s case in Italy suggests that areas which did not develop social capital in the Middle Ages are not likely to perform well in the twentieth century. He seems skeptical that government policy can work to create this trust. We are left then with the possible policy conclusion that decentralization will work only in areas with strong histories of social capital and that the rest of the country should be centralized — a conclusion that is not likely to be politically viable. Nevertheless, the social capital approach does suggest elements of the local context may affect the functioning and effectiveness of decentralization and that studies of decentralization should take this local context into account.

**Principal agent approach**

This approach has also been developed by economists and has been used primarily to examine choices made by managers of private corporations (Pratt and Zeckhauser, 1991). It has also been used by economists and political scientists to analyze federal intergovernmental transfers to states in the United States (Chubb, 1985; Hedge et al., 1991; Frank and Gaynor, 1993). In Britain, it has been used to analyze local governments as agents of the central government (Griffith, 1966) and to examine the bargaining between these levels of government (Rhodes, 1986). In recent years, the principal agent approach has also been used by sociologists, economists and others in the field of health care to analyze the relationship between provider and patient (Dranove and White, 1987).

This approach proposes a principal (individual or institution) with specific objectives and agents who are needed to implement activities to achieve those objectives. These agents, while they may share some of the principal’s objectives, also have other (usually self-regarding) interests, such as increasing their own income or reducing the time and effort they devote to tasks for the principal. Agents also have more information about what they are doing than does the principal, giving them an advantage which could allow them to pursue their own interests at the expense of those of the principal. The principal might like to overcome this information asymmetry, but gaining information has significant costs and may be impossible. So the principal seeks to achieve his objectives by shaping incentives for the agent that are in line with the agent’s own self-interests. The principal can also use selective monitoring and punishments to encourage agents to implement activities to achieve these objectives. In most studies using the principal agent approach, it is assumed that the principal receives the benefits of any profit that is produced by the agents. In addition to the information asymmetry, the principal agent approach also focuses on who controls information and how to improve monitoring (Chai, 1995; Hurley et al., 1995).

This approach allows us to view the Ministry of Health as a principal with the objectives of equity, efficiency, quality and financial soundness (rather than profit as assumed in the economic models). The local authorities are agents who are given resources to implement general policies to achieve these objectives. This approach encourages us to examine how the principal monitors performance and shapes incentives and punishments.

The principal agent approach has advantages over the other approaches reviewed here for developing a systematic framework for research on the decentralization of health systems in developing countries. In contrast to the local fiscal choice approach, which focuses only on the dynamics at the local level, the principal agent approach forces us to look at the relationship between the center and periphery and to see the relationship as dynamic and evolving. The approach, by focusing on the mechanisms that the center can use to shape
choices at the periphery, is also appropriate for providing policy advice to authorities at the national level. It allows us to focus on defining what the national level can do to encourage local authorities to achieve the broad goals of health policy.

Weaknesses often cited are that the principal agent approach focuses on the vertical relationship between the principal and the agent, making it difficult to analyze multiple principals, especially if they are of different administrative levels. Some analysts have taken this problem as a crucial weakness in the principal agent approach (Hedge et al., 1991). Decentralization, at least in its devolution form, implies that those who manage the government bureaucracy will be accountable to the local population (or local political system), who become additional principals and who may have quite different objectives from those of the principals at the national level.

However, the principal agent approach can accommodate multiple principals. While the usual multiple agent analysis has focused on a vertical chain of principals — the “people” as principal who elect the Congress as agent, which in turn acts as principal over the government bureaucracy which acts as agent (Chubb, 1985; Moe, 1991) — multiple principals can be competitive (as in Congress vs the President) and the approach can still inform us on this relationship. There is no inherent logic in the principal agent approach which prevents this analysis from including multiple principals at either the national or the local level.

Nevertheless, when it is applied to the analysis of decentralization, the principal agent approach does have a specific blind spot. It does not have an easy conceptual means of defining the range of choice that is by law and regulation transferred from one authority (the principal) to another (the agent). As it has been applied in the literature, the principal agent approach can be used to analyze both centralized and decentralized systems. The agents in a centralized bureaucracy are subject to a principal’s control through incentives and sanctions and through monitoring, although the types of incentives and monitoring may be different from those in a decentralized system. What is needed to make the approach applicable to an analysis of the effects of decentralization is a means of describing the shift in the range of control that the principal can exercise over the agent. We will return to this point later as we develop the concept of decision space.

**Toward a framework for the study of decentralization of health systems in developing countries**

Each approach we have reviewed has some validity and provides some insight into key issues of decentralization. The public administration approach provides an institutional framework that focuses on types of institutional arrangements. It is useful for describing transfers of authority to different types of institutions (devolution, delegation and privatization). In these cases, it is particularly important to analyze the capacity of the institutions receiving the new powers and authority to take on the tasks assigned. However, this approach, although it is in wide currency now, is not very useful as a framework for analyzing the types of choices made by local authorities. Local fiscal choice is especially useful in focusing attention on the accountability of local officials to local populations (voters/taxpayers). Since it uses assumptions of public choice models, it also proposes a clear set of objectives and/or motivations for generating hypotheses about choices at this level. However, the importance of intergovernmental transfers compared to local funding sources and the restrictions on their use by central governments, limit flexibility and accountability at the local levels, undermining the utility of this approach as a general framework.

The social capital approach suggests that some characteristics of the local community may facilitate the capacity of local governments to perform better and to achieve objectives such as those of health reform. It is a relatively conservative vision, however, that does not have clear policy implications, at least in the form presented by Putnam.

This review suggests that the principal agent framework is likely to be the most effective overall approach to decentralization and that other approaches may offer supplementary concepts and hypotheses. The principal agent framework focuses our attention on the relationship between the center and the periphery and can generate policy recommendations about how the center can shape decisions made at the periphery so that they are more likely to achieve the objectives of health reform. Its major weakness is that it does not have a clear means of defining the range of choice allowed by decentralization. This is the issue we address next.

**MODIFYING THE PRINCIPAL AGENT APPROACH TO ADDRESS DECENTRALIZATION AND HEALTH REFORM: THE DECISION SPACE APPROACH**

The following sections tailor the principal agent approach to the issues of decentralization and the achievement of health reform objectives. The principal agent approach places the issue of decentralization in the context of the objectives of the principal and how the principal uses various mechanisms of control to assure that the agents work toward achieving those objectives. The literature on the principal agent approach identifies several channels of control which are available to the principal. They include: positive incentives, sanctions and information to monitor compliance. I discuss these channels below; however, decentralization requires additional concepts to capture the widening range of discretion or choice allowed to agents in the process of decentralization which differentiates decentralized principal agent relationships from...
centralized relationships. I call this concept “decision space”.

**Decision space**

Decentralization inherently implies the expansion of choice at the local level. We need to develop a way of describing this expansion. I propose the concept of “decision space” as the range of effective choice that is allowed by the central authorities (the principal) to be utilized by local authorities (the agents)*. This space can be formally defined by laws and regulations (and national court court decisions). This space defines the specific “rules of the game” for decentralized agents. The actual (or “informal”) decision space may also be defined by lack of enforcement of these formal definitions that allows lower level officials at each level to “bend the rules”. Decision space may be an area of negotiation and friction between levels, with local authorities often challenging the degree of decision space conferred on them by the central authorities.

Decision space is defined for various functions and activities over which local authorities will have increased choice. It can be displayed as a map of functions and degrees of choice as presented above†. In Table 1, the map of decision space displays (across the vertical axis) a series of functional areas where expanded choice can occur and (across the horizontal axis) an estimate of the range of choice or discretion, (for illustrative purposes defined here as “narrow”, “moderate” and “wide”), that is allowed for that dimension‡ (see Table 2). This approach allows us to disaggregate the functions over which local officials have a defined range of discretion, rather than seeing decentralization as a single transfer of a block of authority and responsibility§.

Table 1. Map of decision space

<table>
<thead>
<tr>
<th>Functions</th>
<th>Range of choice</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>narrow</td>
</tr>
<tr>
<td><strong>Finance</strong></td>
<td></td>
</tr>
<tr>
<td>Sources of revenue</td>
<td>→</td>
</tr>
<tr>
<td>Allocation of expenditures</td>
<td>→</td>
</tr>
<tr>
<td>Income from fees and contracts</td>
<td>→</td>
</tr>
<tr>
<td><strong>Service organization</strong></td>
<td></td>
</tr>
<tr>
<td>Hospital autonomy</td>
<td>→</td>
</tr>
<tr>
<td>Insurance plans</td>
<td>→</td>
</tr>
<tr>
<td>Payment mechanisms</td>
<td>→</td>
</tr>
<tr>
<td>Contracts with private providers</td>
<td>→</td>
</tr>
<tr>
<td>Required programs/norms</td>
<td>→</td>
</tr>
<tr>
<td><strong>Human resources</strong></td>
<td></td>
</tr>
<tr>
<td>Salaries</td>
<td>→</td>
</tr>
<tr>
<td>Contracts</td>
<td>→</td>
</tr>
<tr>
<td>Civil service</td>
<td>→</td>
</tr>
<tr>
<td><strong>Access rules</strong></td>
<td></td>
</tr>
<tr>
<td>Targeting</td>
<td>→</td>
</tr>
<tr>
<td><strong>Governance rules</strong></td>
<td></td>
</tr>
<tr>
<td>Facility boards</td>
<td>→</td>
</tr>
<tr>
<td>Health offices</td>
<td>→</td>
</tr>
<tr>
<td>Community participation</td>
<td>→</td>
</tr>
</tbody>
</table>

*This concept draws on the public administration concept of discretion, which normally distinguishes between political and technical choice and the role of allowing administrators choice within parameters set by legislation. Here, alternatively, I have tried to specify the degree of discretion allowed for specific functions with high technocratic content. On the concept of discretion see: Shumavon and Hibbeln (1986) and Bryner (1987).

†The map matrix presented here is derived from a matrix on hospital autonomy developed by Chawla and Berman (1996).

‡In order to make this map a tool for rigorous comparisons it will be necessary to develop quantifiable indicators for each function. See Table 2.

§The discussion here uses local government (provinces or municipalities) as an example of the local authority that is receiving greater decision space in the process of decentralization. This choice is for simplicity of illustration. The approach could also be used for deconcentrated authority to regional or district offices within the Ministry of Health. It is useful to note also that regional or district offices of the Ministry of Health may share control over the local government with the central authority. In these cases a separate analysis would be necessary to identify the functions and extent of control that the two different authorities exercise over the local government.
Table 2. Indicators for mapping decision space. Below is a suggestive table of indicators that could be examined for comparative mapping of decision space

<table>
<thead>
<tr>
<th>Function</th>
<th>Indicator</th>
<th>Range of choice</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>narrow</td>
<td>moderate</td>
</tr>
<tr>
<td>Finance</td>
<td>Sources of revenue</td>
<td>high %</td>
</tr>
<tr>
<td></td>
<td>Intergovernmental transfers as % of total local health spending</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Allocation of expenditures</td>
<td>high %</td>
</tr>
<tr>
<td></td>
<td>% of local spending that is explicitly earmarked by higher authorities</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fees</td>
<td>no choice or narrow range</td>
</tr>
<tr>
<td></td>
<td>Contracts</td>
<td>none or one</td>
</tr>
<tr>
<td>Service organization</td>
<td>Hospital autonomy</td>
<td>defined by law or higher authority</td>
</tr>
<tr>
<td></td>
<td>Insurance plans</td>
<td>defined by law or higher authority</td>
</tr>
<tr>
<td></td>
<td>Payment mechanisms</td>
<td>defined by law or higher authority</td>
</tr>
<tr>
<td></td>
<td>Required programs</td>
<td>rigid norms</td>
</tr>
<tr>
<td>Human resources</td>
<td>Salaries</td>
<td>defined by law or higher authority</td>
</tr>
<tr>
<td></td>
<td>Contract</td>
<td>none or defined by higher authority</td>
</tr>
<tr>
<td></td>
<td>Civil service</td>
<td>national civil service</td>
</tr>
<tr>
<td>Access rules</td>
<td>Targeting</td>
<td>law or defined by higher authority</td>
</tr>
<tr>
<td>Governance rules</td>
<td>Facility boards</td>
<td>defined priority populations</td>
</tr>
<tr>
<td></td>
<td>District offices</td>
<td>size and composition of boards</td>
</tr>
<tr>
<td></td>
<td>Community participation</td>
<td>size, number, composition, and role of community participation</td>
</tr>
</tbody>
</table>
quality and financial soundness. Key decisions on sources of revenue and allocation of expenditures are likely to have significant influence on equity and financial soundness, although some allocation decisions — for instance, those related to funding for prevention and promotion — may also affect efficiency and the quality of services. Decisions about the organizational structure of services are also likely to have an important impact on efficiency, quality and equity. Allowing competition among providers and insurance plans and between public and private entities may increase efficiency and quality of service. Increasing flexibility on decisions about human resources — particularly allowing for productivity and quality incentives for providers and allowing managers greater ability to hire and fire — may increase efficiency and quality of services. Restricting access to facilities or eligibility for subsidies is a classic tool for achieving equity objectives by allowing scarce public resources to be targeted to the poor.

Finally, governance rules influence the roles local political actors, beneficiaries and providers can play in making local decisions. These rules structure local participation in a decentralized system*.

For example, in Colombia, where devolution to departments (similar to provinces or states) has been implemented over the last five years, the matrix in Table 3 could be used to define the formal range of choice in five major functional areas allowed to local authorities. It specifies choice that is defined by a series of laws and regulations through which the central government devolved power to the departments.

This map shows that for finance functions the decentralization process in Colombia has allowed local authorities a moderate range of choice over sources of revenue from intergovernmental transfers (by a formula which assigns a minimum percentage that must be assigned to health and a percentage over which local discretion is allowed)†. Some local revenues (taxes on liquor, beer, tobacco and lotteries) are assigned to secondary and tertiary health facilities by law. Other local revenues (which average only 10% of total local revenues) can be assigned to health at the complete discretion of the department government.

For decisions on allocating expenditures, the range of choice for the departments is moderate. The department government is directed to assign 50% of one source of intergovernmental transfer (the situado fiscal) to primary health care, transferring it to the municipalities that operate the primary level facilities. Of the remainder, 40% must be assigned to the secondary and tertiary care facilities and 10% must be assigned to a basic public health benefits package (the Plan de Atención Básica — PAB). The fee structure of hospitals in Colombia is determined by the hospital board so the department government only has a role as participant in the board’s decisions.

For Colombia’s departments, the decision space for a service organization is generally quite wide. While hospital autonomy is defined by law to hospitals are supposed to have strictly defined tripartite boards with fairly wide powers — under current law departments are allowed a range of choice on

*Of course, with multiple objectives some activities may support some objectives at the expense of others. The framework here could be used to evaluate these results and to suggest means of maximizing the objectives.
†The original legislation (Law 60) “forced” the departments to assign 60% of the situado fiscal to education, 20% to health and the remaining 20% could be assigned at the discretion of the department to either health or education. This flexibility has been reduced by a recent law which removed the discretion over the “unforced” percentage.
how to contract with insurance plans. The departments themselves can act as public insurance providers (at least until significant private competition is available), they can contract with special publicly designed insurance plans, or they can contract with private plans. Norms and standards of Colombian national health programs are quite restrictive in some areas — for instance, in defining staffing patterns and architectural requirements for hospitals, — but in other areas, such as quality and coverage objectives, the standards are not well defined.

The mechanisms that the departments in Colombia use to pay providers are also open to a wide range of options, from supply side subsidies to hospitals, to fee for service, to per capita payments and mixed payment schemes.

In the functional area of human resources, salary levels for permanent staff are defined in Colombia by a national salary leveling law. These levels appear to be the floor for salaries and some discretion is allowed to local authorities to “top up” salaries. Contracts for non-permanent staff are not specifically restricted by law or regulation. Hiring and firing of permanent staff, however, is severely restricted by civil service laws that apply to all personnel. Contracts for non-permanent staff are not specifically restricted by law or regulation. Hiring and firing of permanent staff, however, is severely restricted by civil service laws that apply to all personnel.

National laws in Colombia also strictly define who is eligible for access to subsidized facilities and health plans. The targeting mechanism is a nationally designed census that identifies socio-economic strata by family (SISBEN). Local governments are required to implement this census and to distribute identity cards to the families. Governance rules for hospital boards, local offices and arenas for community participation are also defined by law.

It is important to note that this formal map of the decision space may not reflect the actual range of decision available to local authorities. The formal laws and regulations may not be enforced and may be violated either by the agent or the principal. The agent may make decisions that are not formally allowed and the principal may in practice restrict choice that is formally allowed to the agent. In such a case, it would be necessary to develop an “informal map of decision space” to identify whether legal and regulatory rules have been respected or whether the actual range of choice is different. In Colombia, for example, many municipalities which are legally certified to exercise wide discretion are still centrally controlled in some of the functional areas, while other municipalities, which are not formally certified, are able to exercise decisions in functional areas for which they have no legal decision space (Jaramillo, 1996).

Viewed from the perspective of the agents, the decision space is a channel of central control. It is one of the mechanisms the center uses to try to get the agents to achieve the center’s goals. At the center, however, the decision space is the product of a variety of decisions made by various actors and in this sense it may be a channel of control of multiple principals in the center. The decision space may be partly defined by legislation in which both the Ministry of Health and the decentralized units are bound. The ministry’s ability to change the decision space and even to provide incentives and punishments is limited by decisions made by the other institutions of the central government. For instance, in Colombia the ministry cannot change the general rules for allocating revenues to the departments without proposing major changes in the laws. However, the ministry can change the regulations on competitive bidding for insurance plans for the subsidized population, opening new options for insuring this population.

In the following discussion our focus will be on analyzing the ministry as principal and the local health authorities as agents; however it is important to keep in mind the restrictions that are placed even on the ministry by other principals in the center.

Use of decision space: innovations, directed change and no change

The second set of unique questions that decentralization raises is the response of the agent to the discretion allowed by a wider decision space. The agents who are allowed wider discretion may choose not to take advantage of the new powers and simply continue to pursue activities as they had before. Alternatively, they may choose to innovate by making new choices they had not made before. Innovation has become a central issue of investigation for programs promoting local government in the U.S. (Altshuler and Behn, 1997) and in Latin America (Campbell, 1997). Innovation can be seen as having three dimensions, temporal, functional and structural. Decentralized authorities innovate in a temporal sense when they make decisions that are different from those they made before decentralization. Local agents may also innovate in one or more functional area and not in the others for which they have wider discretion. Finally, the localities that enjoy a relatively wider range of choice in their decision space innovate when they make decisions that are not available to localities that are controlled by central decisions.

Centrally controlled localities may also make what we might call “directed change”. The central authorities may promote significant directed changes over time — changes that non-decentralized localities are forced to adopt but the decentralized authorities are not required to make. In these cases the non-decentralized units are changing policy and the decentralized units are not. If the decision space is characterized by a wide range of choice but local officials simply continue to do what they had been doing under the centralized system, then a wider decision space has not resulted in innovative local choice.
The use of decision space might be analyzed along the functional dimensions of the map of decision space above to see: (1) whether or not changes were made, (2) in cases where there were changes, whether or not they were innovations or just directed changes and (3) how these innovations or directed changes affect the performance of the local health systems in achieving health reform objectives.

Performance

Next we need to determine which of the choices — innovations, directed change, or no change — is likely to achieve the objectives of health reform. We will need to determine whether the wider decision space and the capacity to innovate, to reject “directed change”, or simply to continue doing what was done before, is likely to improve the capacity of a nation to reach its health reform goals. Therefore it becomes essential that we evaluate the “innovations”, “directed change” and “no change” in terms of their impact on performance in areas defined by the objectives of health reform.

Much of the argument over different policy choices at any level of government is an argument about the likelihood of different mechanisms, tools and institutional arrangements to achieve the broader objectives of a health system. There is no clear evidence to suggest that we know what combined package of policies can maximize the achievement of the objectives of equity, efficiency, quality and financial soundness. Both central governments and local governments can make choices of policies that might or might not achieve the objectives. Some choices may lead to achievement of one objective at the expense of others. Furthermore, many of these objectives are also influenced by other factors that are outside the control of either level of government. We therefore must enter this territory with some caution. However, it is through measures of performance that we can establish whether and by what ranges of decision space, decentralization can assist a country to achieve the objectives of health reform.

There are some choices which we have some reason to believe are effective in reaching health reform objectives, either by strong theoretical logic or experience in other countries. There are other choices whose effectiveness is less well understood. Current thinking suggests that separating financing and provision of service (for instance by introducing insurance plans between the financing and the providing institutions) and introducing some level of competition is likely to improve efficiency of health services and might also improve quality (World Bank, 1993). We also have some evidence that the ability of local managers to hire, fire and provide specific incentives to employees improves efficiency (Chai, 1995). We assume often that increased funding for health is likely to improve quality and, if targeted correctly, improve equity.

However, evaluating performance is a significant task. The central problem with the evaluation of performance is the lack of reliable data on all dimensions of the overall objectives. Recent examples of indicators of performance which have been used in studies of decentralization tend to focus on expenditures. Per capita spending is used as an indicator of equity (Putnam, 1993; Carciofi et al., 1996; Jacobsen and BeGuire, 1996). Other studies have examined the decline in local counterpart funding generated by a growth in intergovernmental transfers as an indicator of “fiscal laziness” or lack of assumption of fiscal responsibility by local authorities (World Bank, 1994; Kure, 1995; Wisner Duran, 1995). Putnam has also used an index of general performance to evaluate decentralized institutions in Italy. This index uses measures from all sectors, including only two from the health sector: number of family clinics and local health unit expenditures per capita.

The following list suggests some potential indicators of performance:

- Equity
  - changes in coverage by insurance programs
  - changes in per capita spending
  - changes in local vs national revenue sources
  - percentage of targeted population subscribed in insurance plans
  - changes in utilization by socio-economic strata

- Efficiency
  - changes in hospital productivity
  - changes in bed occupancy rates and lengths of stay

- Quality
  - changes in intra-hospital infection rates
  - changes in immunization coverage and low birth weight
  - changes in patient satisfaction

- Financial Soundness
  - funding/subsidized regime
  - hospital deficits

Studies will have to develop these indicators based on the availability of reliable data.

Positive incentives and sanctions

The principal does not rely only on the formal “decision space” to encourage local agents to achieve the objectives of health reform. Other channels of control used by the principal are the rewards and punishments that the principal can use to entice the agents to achieve the principal’s objectives.

Incentives may be defined in both individual and institutional terms. The incentives of intergovernmental transfers usually are defined in terms of in-
stitutions, since the entity receiving the funds may be the municipal or provincial government. However, it may also be important to evaluate the individual incentives of major decision-makers within these institutions.

The flow of additional resources as intergovernmental transfers might be seen as an incentive to the local authorities, especially if these resources can be taken away by the principal if the locality does not achieve objectives or follow administrative rules. One particularly important perverse incentive is the granting of discretionary funds to cover deficit spending — the “soft-budget” constraint. Other mechanisms of incentives might be the achievement of benchmark targets which trigger additional funding, or different ratios of matching grants (Frank and Gaynor, 1993). In some cases, the granting of wide decision space is an incentive in and of itself. This is an important incentive for professionals within organizations and may be an incentive for local institutions. The following list of incentives is an example of possible incentives that the principal can offer:

Economic incentives

- manipulating the formula for the allocation of intergovernmental transfers to departments and municipalities to reward the agents who achieve specified objectives
- Ministry of Health discretionary funding for investments, covering deficits and other operating costs — through control of some discretionary ministry budgets, through social investment funds and through influence over donor funding
- manipulation of matching grant requirements for local resources

Non-economic incentives to individual officials

- Fellowships
- Career advancement
- Opportunities for corruption

Non-economic incentives to departments and municipalities

- Technical assistance
- Wider decision space

Economic incentives to individual officials

- Wider decision space
- Professional training
- Recognition for achievement

Sanctions might include reduction of transfers for failure to achieve objectives, intervention or takeover by the center for flagrant disregard for rules and regulations or failure to provide minimal health services. Sanctions include withdrawal of any of the positive incentives above and:

- Fines and jail (for breaking rules of formal decision space)
- Intervention (takeover by higher authorities)
- Firing officials

Incentives and sanctions are central issues within the principal agent approach. A wealth of potential hypotheses about incentives and sanctions has come from the theoretical and empirical work that has been done to date. Much of the literature about principal agent relationships revolves around how the principal can set incentives so that agents have a stake in achieving the principal’s objectives. Not only the type and level of incentives are seen as important, but also the structure providing the rewards and sanctions is crucial.

Information and monitoring

Information and monitoring are crucial for the principal to evaluate how and whether the agents are achieving the principal’s objectives. But information and monitoring have significant costs. However, the agent’s control of information is crucial to the negotiating power of the agent vis-à-vis the principal.

Central ministries often have routine information systems through which their agents must report. The information available to the principal is usually of variable quality and can often be manipulated — through failure to report or through inaccurate reporting — by the agent. This information often includes utilization, coverage, human resources and budgets. Budgetary categories are usually not designed for assessing achievement of health reform objectives. It is therefore important to assess how much information is available to the central authorities, the capacity of the central authorities to process this information and the quality of the information.

CHARACTERISTICS OF THE AGENT

The characteristics of the agent will also influence how it responds to the mechanisms of control and how it pursues innovations. These characteristics can be classified as being related to (1) the motivations and goals of the agents, (2) the role and influence of local principals and (3) the capacity of the local agents to innovate and implement.

Motivations and goals of agents

Some of the literature on the principal agent approach suggests that if the goals and motivations of both the principal and agent are compatible, then the principal-agent relationship will be more effective (Pratt and Zeckhauser, 1991). The central assumption of most principal agent literature is that agents (as individuals and, by extension, institutions) are self-interested and concerned mainly about maximizing control of finance and leisure. If these assumptions are correct, all agents will have these motivations and incentives will have to be directed toward achieving them. While these assump-
tions assist in the formulation of theory and hypothesis, they do not always explain actual behavior. Several other motivations are discussed in the literature: professional approbation (Wilson, 1989; Eisner and Meier, 1990), achievement of a specific institutional mission (Bullock and Lamb, 1984; Weiss, 1996) and organizational survival (March and Simon, 1993). These motivations should be examined in relation to the objectives of and incentives offered by the principal.

Local principals

In decentralization cases where there has been an institutional break — as in devolution, delegation or privatization — it is likely that some form of multiple agency analysis would be necessary to appraise the results of decentralization. Since the health authorities in local governments must respond in part to elected officials (mayors, governors, legislators), who in turn are agents of the principals in the local political process (electorate and/or dominant political coalition), the goals and interests of these local principals will shape the response of the municipal health officials to the incentives and rewards of the central government.

The role of the local political process can be examined by a variety of methods, from stakeholder analysis to median-voter public choice models. An initial study might focus on a stakeholder analysis of the local municipality or province, examining the power of different local interest groups, especially the power of physicians, insurance companies and hospitals. As Wilson (1989), points out, those interests which are concentrated and have significant investments are likely to have more influence over bureaucracies than are the dispersed beneficiaries who have only sporadic interest in health issues. The extensive literature on interest group politics in health care could provide additional hypotheses for local level decision-making (Eckstein, 1958; Marmor, 1973; Reich, 1995). It would be particularly important to examine the mechanisms used for community participation to balance out the influence of the vested interest groups. Here again, there is a wide literature on community participation and local accountability from which to draw hypotheses (Esmann and Uphoff, 1984; Paul, 1992).

Once the objectives of the local principals are defined by this kind of analysis, we would have to analyze the range of incentives and sanctions that these principals can exercise over the local health administrators. These incentives and sanctions, which can complement or undermine those of the central principals, can be related to the local capacity to mobilize its own funds, the capacity to hire and fire administrators, or the opportunities for professional recognition or corruption. A major incentive of these local principals will be in the provision of additional local funding. If the local political process allows significant contributions from local own-source revenues (which are not already earmarked by the decision space), then the dynamics suggested by the local fiscal choice literature may be useful to examine. In any case, local principals with considerable additional resources are likely to have greater influence vis-a-vis the principals in the “center” and the conflict in objectives may become more pronounced. An alternative situation may be one in which local resources allow local principals to dictate particular innovations that are not available to centrally directed localities or to poorer localities without sufficient additional resources to assign. We address this latter issue below.

Capabilities of the agent

The capabilities of the agents may also be an important set of variables defining the agents’ response to the principal. Of the characteristics that might influence the capacity of agents to make decisions that are likely to be responsive to the objectives of the principal, we focus here on the issues of human resource capabilities, socio-economic characteristics and social capital.

First, the human resources available in the municipal or province may condition the ability of the agent to make decisions within the decision space allowed. Communities with few professionals or those with the wrong professional mix, may not perform as well as others with a similar decision space. There are some studies on the relationship of technical capability to organizational performance which can be used to develop hypotheses on this issue (Scott, 1987). It may also be interesting to compare locally recruited professionals to those who are recruited through the national centralized system. In some cases, the staff in the newly decentralized unit were simply transferred from the local staff of the ministry. This is the case in many deconcentrated forms of decentralization and may also be the case in devolved forms where the regional staff is simply moved from the regional office of the Ministry of Health to the provincial governor’s office. This transfer may bring appropriate skills that would be lacking in newly created offices, but it also retains the structure, culture and routines of a highly centralized institution.

Socio-economic characteristics of the local municipality or province might also affect the capacity of the agent to implement innovations. Those communities with a larger local resource base may be able to assign local resources to complement those of the intergovernmental transfers. Higher socio-economic status may also bring a larger pool of trained personnel and other advantages which strengthen its capacity to implement what the principal desires. However, wealthier communities also may have more political power in the national political process and can refuse to accept the directions, incen-
tives and sanctions that the Ministry of Health might impose.

Using Putnam’s analysis of social capital, we might hypothesize that communities with denser networks of civic organizations will have greater social capital that will strengthen their capability to choose innovations and implement health programs effectively. Alternatively, Putnam suggests areas without social capital may be dominated by clientalism, which is based on vertical relationships of “instrumental friendships”, which are less likely to be able to perform well. Putnam has used surveys to identify the areas where community members feel that clientalism is strongest as one indicator of the lack of social capital.

POLICY IMPLICATIONS OF THE DECISION SPACE APPROACH

The decision space approach has some direct implications for policy choice at the central level as well as at the local government level.

National governments generally control the rules and processes of decentralization. We want to be able to advise them on how to shape the decision space for local governments so that these governments will make choices that are more likely to achieve the desired levels of performance. The decision space approach will give us some idea of how much discretion, over what kinds of functions, is likely to lead local authorities to make choices that will achieve central government objectives. How wide should the “decision space” be so that the central resources are used appropriately, or matched appropriately by locally generated revenues? Should discretion over local salary levels be wide or narrow to improve the efficiency of the work force? Should local authorities be allowed to choose different mechanisms for dividing financing and provision, or should they be forced to make only one or two choices? What choices can maximize the achievement of all health reform objectives and what choices favor some objectives at the expense of others?

We could then examine the role of incentives and sanctions used by the Ministry of Health to shape the choices of local health administrators. How have intergovernmental transfers been used to achieve the goals of health reform? What has been the role of perverse incentives for deficit spending? Have there been cases of central intervention of local health administrations? How have career and professional incentives been used? With preliminary answers to these questions we could then recommend improved use of these incentives and sanctions, tailoring them to achieve indicators of performance.

The decision space approach also might suggest that some choices should be limited to local govern-
ments with specific characteristics — for instance, only governments which are likely to have sufficient administrative capacities. It may also be important to assure that local participation mechanisms are in place for the poor to have a voice in decision making.

The decision space approach, drawing on a principal agent analysis, also suggests that the center should focus on developing appropriate information systems and indicators so it can efficiently monitor the behavior of local governments and apply sanctions or incentives appropriately. Lack of key data may inhibit the ability of the ministry to get local governments to achieve its objectives.

In order to address these policy issues our central research questions that emerge from the decision space approach can be summarized as:

- What is the effect of larger decision space on the taking of innovative decisions?
- What is the effect of centrally controlled incentives and sanctions on the choices of local health administrations?
- What explains why some local health administrations implement innovations and others do not?
- Do these innovative decisions make a difference in performance?

Acknowledgements—This article is based on research supported by the Data for Decision Making Project (funded by United States Agency for International Development, Cooperative Agreement DPE-5991-A-00-1052-00) and the Colombia Health Sector Reform Project (funded by the Government of Colombia and the Interamerican Development Bank, Contract No. 001/95) both based at Harvard University School of Public Health. The author wishes to thank the following colleagues for significant comment and suggestions on drafts and presentations at a seminar at the Ministry of Public Health in Colombia and the Health Systems Studies Seminar of the Program in Health Care Financing, Harvard School of Public Health: William Hsiao, Merilee Grindle, Michael Reich, Peter Berman, Jack Needleman, Winnie Yip, Mukesh Chawla, Kara Hanson, Jorge Enrique Vargas, Deyana Acosta-Madiedo, Antonio Mendoza, Gonzalo Leal, Bernardo Barona, Fernando Rojas, Jacqueline Arzoz, Lida Alacón, Vivian Goldman and Claudia Reeder. He also thanks Anthony Zwi and the two anonymous referees. The author is solely responsible for this article.

REFERENCES


