

Intergovernmental Fiscal Relations in Latin America:
Policy Design and Policy Outcomes

Richard Bird

Inter-American Development Bank

Washington, D.C.

Sustainable Development Department

Foreword

The Sustainable Development Department commissioned Richard Bird to prepare a paper on intergovernmental relationships as part of the activities leading to the preparation of a Subnational Development Strategy for the IDB. Using his vast academic and consulting experience, Bird produced a document of significant breath and depth, touching on the most important aspects that determine the complex relationship that exist between central and subnational government, with a particular reference to Latin America and the Caribbean. His analysis provided key inputs to the work of the team drafting the strategy paper and portions of it were included in the background paper for the strategy. However, given the many insights that this work provides on issues surrounding intergovernment relations, the Social Development Division considered worth publishing the study in its entirety.

The paper develops recommendations concerning a desirable structure of intergovernmental relations that create the correct incentives for subnational governments to become efficient and equitable providers of services. However, as Bird highlights in his conclusions, the discussion and recommendations contained in the paper, are of a general nature and point towards an optimal structure of intergovernmental relations. They need to be adapted to the realities of each individual country and gain their full significance when full consideration is given to the legal, political and socioeconomic determinants of each individual case.

We are pleased to publish the full text of Bird's paper as a guide for policymakers and practitioners interested in the topic and as a contribution to promote in-depth analysis of this complex area of government.

Mayra Buvinic
Chief, Social Development Division

Contents

Intergovernmental Fiscal Relations	1
The Context of Intergovernmental Fiscal Relations	2
Models of Local Government Finance	
Central Government Policy	
The Big Questions	8
Expenditure Assignment	
Revenue Assignment	
Closing the Gap	
Other Issues	
Government Spending	11
Assigning Expenditures	
Managing Expenditures	
The Revenue Side	16
Tax Assignment in Principle	
Financing Regional Governments	
Financing Local Services	
The Future of Tax Assignment	
Designing Intergovernmental Transfers	25
Key Design Options	
Fiscal Capacity and Fiscal Effort	
Transfers and Poverty	
Conclusion	
Improving Local Capacity	32
Macroeconomic Aspects of Decentralization	35
Transfers and Deficits	
Subnational Borrowing	
Information and Accountability	40
Institutionalizing Decentralization	
Achieving Fiscal Transparency	
Conclusion	45
References	47
Annexes	55

Intergovernmental Fiscal Relations

Decentralization is an important phenomenon in an increasing number of Latin American countries (See IDB, 1994, 1997; Burki et al., 1999; ECLAC/GTZ, 1996, 1997; Lopez Murphy, 1995; Fukasaku and Hausmann, 1998). Subnational finance and intergovernmental fiscal relations have long been matters of concern in Brazil and, to a lesser extent in Argentina. These issues are now becoming increasingly important in other larger countries, whether formally federal, such as Venezuela and Mexico, or unitary, such as Colombia. Even smaller countries such as Paraguay, Uruguay, Ecuador, and Guatemala are increasingly considering decentralization. Decentralization is not only in the air, it is already on the ground and influencing policy outcomes throughout the continent in a major way. It is therefore critical to understand the actual and potential impact of intergovernmental relationships on such key questions as the ability of central governments to conduct sound macroeconomic policies, the incentives for subnational governments to behave in a fiscally responsible way, and the efficient and accountable delivery of public services.

This paper sets out some of the main factors determining the effects of intergovernmental fiscal relations on such matters. It represents not new research but rather an attempt to digest and reflect upon what we have learned to date about some of these complex questions.¹ The balance of the paper is organized as follows. First, since the structure and to some extent the effects of intergovernmental fiscal relations depend largely upon

the objectives and constraints of the country in question, Section 2 discusses some core general questions that arise with respect to decentralization issues — specifically, the objectives of decentralization, the related but distinct question of the underlying model of subnational finance, and the key role played by the central government in decentralization. Against this background, Section 3 briefly introduces the three key components of intergovernmental fiscal relations — expenditures, revenues, and transfers. Sections 4 through 6 review, under each of these headings, the implications of alternative arrangements for the efficient delivery of public services, emphasizing the importance of both enabling and requiring hard budget constraints at all levels of government. Some have argued that the outcome of recent decentralization in many countries will be unsatisfactory owing to local incapacity to handle the new tasks. Others have asserted that decentralization will result in macroeconomic problems. Section 7 considers briefly the question of local capacity, while Section 8 discusses subnational borrowing and the implications of decentralization for macroeconomic management. Section 9 sets out the critical need to develop both an adequate information base and an appropriate institutional forum if the dynamics of decentralization are to evolve in a constructive direction in any country. Section 10 summarizes some of the key conclusions of the paper.

¹Much of the paper draws freely on related prior work. See, for example, on Colombia (Bird, 1984; Bird and Fiszbein, 1998; and World Bank, 1996a) and on Argentina (World Bank, 1996b). I have also made use of additional unpublished material (initially prepared largely for the World Bank at various times) dealing with these countries and with Mexico, Venezuela, Chile, and Uruguay, as well as on more general published materials such as Bird (1993), Bird and Vaillancourt (1998), and Litvack, Ahmad, and Bird (1998). Some of the issues discussed in this paper are treated in more detail in Bird (1999a, 1999b, and 1999c).

The Context of Intergovernmental Fiscal Relations

Intergovernmental fiscal relations are only one aspect of a complex political, economic, and administrative system. Their design, role, and effects can only be understood in the specific institutional context in which they operate. How one views many issues, for example, depends in part upon whether one views decentralization as a process of delegation or devolution. *Delegation* refers to a situation in which local governments act as agents for the central government, executing certain functions on its behalf. The appropriate perspective is “top-down”, that is, from the viewpoint of a central government² whose objective in decentralizing is often assumed, in theory, to be the achievement of a more efficient allocation of resources through the delegation of authority to local governments. The top-down approach implies that the main criterion for evaluating fiscal decentralization should be how well it serves national policy objectives, whatever they may be.³

In contrast, *devolution* refers to a situation in which not only implementation but also the authority to decide what is done is in the hands of local governments — or, to put it another way, local autonomy prevails. The appropriate perspective is then “bottom-up,” from the viewpoint of local citizens. Those who take this approach often stress such political values as improved

governance (through increased local political participation, for example) as well as such economic values as allocative efficiency (through increased responsiveness to local preferences or, perhaps, the increased scope for dynamic innovation that may be afforded by a variety of competing local governments). If this is the relevant perspective, the appropriate criteria for appraising intergovernmental fiscal relations may differ sharply from those under the top-down approach.

For example, is a good outcome one which best achieves the goals of the central government or one which frees local governments most from central dictates? Decentralization is unlikely to produce precisely the expenditure pattern that the central government would itself choose to implement unless, improbably, the goals of central and local governments coincide and all decisionmakers face exactly the same incentives as in a centralized system. Since conflicts between central and local governments as to what should be done are inevitable on many issues, the appropriate choice of perspective is essential in analyzing and interpreting issues of fiscal decentralization in any particular jurisdiction.

Similarly, how one evaluates any particular aspect of intergovernmental relations depends in part upon whether one focuses on its *intrinsic* or its *instrumental* aspects. Devolution may, for example, be considered to be an intrinsically worthwhile objective. Whatever outcomes emerge from a decentralized system of decision-making must, from this perspective, be right. Local people may make wrong decisions from the perspective of the central government or of an outside observer, but if *they* make them, the decisions must, by definition, be assumed to be right for them. The results of a good process must themselves be good.

For such arguments to be persuasive, however, several stringent conditions must be satisfied. First, local decision-making processes must be fully democratic. Electoral democracy alone is

²In some federal countries, the “central” government from the perspective of local governments might be the state or provincial government. Although for the most part the terms “subnational” and “local” are used more or less interchangeably in the present paper, as Bird (1995a) stresses, there may be important differences between regional and local governments, particularly in formally federal countries (see also Sections 5 and 10 below).

³The final qualifying phrase is important, since decentralization in the real world has seldom been driven by efficiency objectives. Even those concerned solely with such objectives must take explicitly into account the inevitably political context of most decentralization: see, for example, Salegh and Tomassi (1999) and Haggard (1999).

not enough to satisfy this criterion. The costs and benefits of decisions must be transparent, and everyone affected must have an equal opportunity to influence the decision. Few, if any, countries can meet this high standard. Second, the costs of local decisions must be fully borne by those who make the decisions. There must be no “tax exporting” and no funding at the margin via transfers from other levels of government. Third, the benefits (like the costs) of local decisions must not “spill over” jurisdictional boundaries without explicit and transparent compensation. In principle, the last two conditions, unlike the first one, can be met to a substantial extent by good institutional design, as discussed later in this paper.

Again in principle, even if decentralization is less than perfect it may yield a more efficient and equitable pattern of public services than the over-centralized and unresponsive public sectors that currently exists in many developing countries. But it will do so in practice only if it is properly implemented. The key to ensuring that incentives are conducive to good decisions is to ensure that those who make the decisions bear the financial (and political) consequences. Thus, the three conditions can be extended and restated as political, economic and administrative requirements.

Political Condition: Leaders at all levels should be responsive and responsible to their constituents, and those constituents should be as fully informed as possible about the consequences of their decisions and those of their leaders. Making politicians bear the consequences of their own mistakes is as close as one can get to a *hard political budget constraint*.

Economic Condition: It should be difficult for local residents to shift costs to nonresidents who do not receive benefits, and local decisionmakers should be appropriately accountable to three distinct, though overlapping, groups: (1) their citizens for the use they make of revenues collected from them (through local taxes), (2) users of local public services for the use made of the revenues they contribute (through user charges of various sorts), and (3) taxpayers in general for the use made of any transfers (or subsidized loans) they receive from the central government.

Administrative Condition: Such a system requires an appropriate institutional framework, including, for example, a clear set of framework laws (on local budgeting, financial reporting, taxation, contracting, dispute settlement, rules to be followed in designing user charges, etc.), as well as adequate institutional support to enable subnational governments to operate effectively in this environment.

It is of course much easier to lay down such general prescriptions than to satisfy them in the very diverse situations found in the real world. Nonetheless, if these conditions are not met, the perverse incentives that too often distort public sector outcomes in many countries may be exacerbated by further decentralization.⁴

The recent wave of decentralization in Latin America has largely followed, and accompanied, an increase in the level of democratization (IDB, 1997). As already emphasized, *democracy* is a particularly important mechanism to promote accountability. At the same time, however, *accountability*—in the sense of transparent decision-making—is an essential ingredient of democracy. The ultimate assurance of good government is, on one hand, the ability of citizens to compare governments in terms of the services they provide and the tax prices they charge and, on the other, the ability of citizens to affect and alter the decisions of government (Breton, 1996). Thus, democracy without good information is not enough.

⁴Along similar lines, Bahl and Linn (1994) noted that “...the situation in a developing country that could provide maximum gains from a more decentralized local government structure would include: (1) enough skilled labor, access to materials, and capital plant to expand public service delivery when desired; (2) an efficient tax administration; (3) a taxing power able to capture significant portions of community income incrementally; (4) an income-elastic demand for public services; (5) popularly elected local officials; and (6) some local discretion in shaping the budget and setting the tax rate.” They might have added: (7) a tradition of local decision-making. All these points seem correct, and are discussed in this paper, albeit with somewhat different emphasis.

Even in countries without well-established democratic institutions, in which decentralization is simply an instrumentality of the central government, good information is essential to ensuring enhanced service outcomes. Regardless of the rationale and circumstances surrounding the decentralization process, its outcomes will improve as more information is made publicly available. The extent to which information relevant to understanding and evaluating the impacts of decentralization is available is thus critical to improved outcomes, almost regardless of the status of political democracy. From this perspective, will be seen below, it is important that intergovernmental fiscal relations in decentralized countries be as *simple, transparent, comprehensible, and predictable* as possible.

When both democracy and good information systems are in place, *devolution* (in the sense defined earlier) makes sense. When they are not, it may not. Even if one or more of these conditions does not hold, however, the *delegation* of implementation responsibilities to local bodies may make instrumental sense provided that the incentives facing local decisionmakers are properly structured. In these circumstances, this means that they should be structured to produce the results desired by the central government. Indeed, whether the underlying policy intention is devolution or delegation, central governments must carefully consider the effects of alternative intergovernmental fiscal relations upon the outcomes in which it is interested. The instrumental character of intergovernmental fiscal relations their *effects* on relevant outcomes, is thus critical viewed from any perspective.

Models of Local Government Finance

Another important dimension of intergovernmental finance is the underlying model of local government finance. Two quite different models implicitly underlie much of the discussion of intergovernmental fiscal relations. The first of these models views subnational government in an ability-to-pay framework. This *ability* model focuses on the effects of governmental action, whether by central or local governments, on the interpersonal distribution of income and wealth. In this frame-

work, decentralization is judged largely by its effects on such distributional goals as poverty alleviation. The *benefit* model, in contrast, views subnational governments primarily as agencies providing identifiable services to identifiable local residents (Break, 1992).

The Ability Model. The ability model dominates public discussion in most countries. However, while attention must be paid to the redistributive aspects of public policy at all levels of government, the allocative costs of distorting local taxing and spending decisions in the name of equity may often exceed any conceivable equity gain. Pursuing active distributive policies at the local finance level, though to some extent perhaps inevitable in the context of democratic politics, is nonetheless generally inadvisable from a strictly economic point of view.⁵

In developing countries, most productive tax bases are invariably taken by the central government. Thus, in practice, the ability approach to local government usually amounts to a system in which central transfers finance most local services even though it is often far from clear that those who benefit from such central largesse are less able to pay for what they get than those whose incomes are reduced by the taxes that finance the transfers. If local governments attempt to implement differential policies (such as progressive taxes or “pro-poor” pricing policies), the result over time may be an outflow of tax base and, consequently, a smaller total package of local services than if the benefit approach had been followed.

Moreover, wealthier local governments, which can apply lower tax rates and obtain similar yields, may attract tax base from poorer ones, thus accentuating initial interjurisdictional disparities. Although intergovernmental fiscal relations

⁵Governments are not simply providers of services. They are also institutions which reconcile and manage conflicting interests in society, and distributional policy is one of the major policy instruments used for this purpose. Even if, as is often the case, subnational attempts at redistribution are vitiated by market forces, the *attempt* to redistribute may still be politically essential. For an extended discussion of this point, see Bird (1980).

should presumably not accentuate market-driven inequalities, it is generally a mistake to design tax and expenditure policy specifically to alleviate such disparities. As discussed further in Section 6, attempts to incorporate regional developmental goals into the design of intergovernmental transfers often result in complex and confused policies that not only fail to achieve these objectives but also fail to foster the efficient and effective delivery of desired local public services (Bird, 1982). Whether “ability” concerns are motivated by considerations of interpersonal distribution or interregional disparities, they are better addressed directly rather than indirectly through distortions in the design of intergovernmental fiscal relations.

The Benefit Model. The benefit model of local finance is more conducive to effective, efficient, and accountable local government. In this model, local governments are essentially viewed as equivalent to firms that provide services. Residents are willing to pay a price or charge for the last (marginal) units of government services that they receive that is just equal to the benefit they derive. This approach to local finance is appealing because the results are not only allocatively efficient but also equitable in the sense that no one pays less (or more) than he or she would be willing to pay in a free market.

The benefit model is not without its problems, however. First, it is often surprisingly difficult to design an appropriate pricing policy for many local public services (Annex 3). Second, even if such prices can be designed, implementing them is seldom politically appealing, particularly if, as in most countries, the introduction of financing via user charges means a sharp change in historic practice. Understandably, citizens seldom appreciate being asked to pay for services that were supplied free of charge in the past. Despite these problems, in principle, services that flow to identifiable individuals (or firms) should be priced appropriately to the extent feasible.

Where such pricing is not possible or desirable, local expenditures and revenues should be linked through matching service benefit areas to the spatial dimension of the financing sources. Any taxes levied by subnational governments to fi-

nance local services should thus fall primarily on local residents. Moreover, as discussed further in Section 5, if such governments are to behave in a fiscally responsible way, they must be made politically responsible for determining the rates of at least some major taxes. Subnational governments may have large revenues from what appear to be local taxes, but if they cannot set tax rates they cannot be made accountable to their constituents at the margin, as both democracy and efficiency require.

Finally, even if local services are largely paid for through locally-determined (and locally-paid) fees and taxes, intergovernmental transfers must still be carefully designed to ensure that, *at the margin*, the costs and benefits of local fiscal decisions are borne locally, while taking adequately into account such interjurisdictional spillovers as are deemed relevant. Consistent application of these rules will impose what is called a *hard budget constraint* on local decisionmakers and, hence, make them fully accountable for the consequences of their decisions.⁶

Central Government Policy

The institutional setting within which local governments in most developing countries must function was recently characterized as falling into one of three categories: (1) the over-controlled local public sector, (2) the under-controlled local public sector, and (3) the perversely regulated local public sector (World Bank, 1995a). The first of these situations seems most common in developing countries. Central governments sometimes control all the details of local government (who they hire, what they pay, where and when the buses run, etc.) and leave little or no freedom for local initiative.⁷ Central governments also often finance

⁶Of course, neither local decisionmakers nor, in most instances, their constituents may be happy to be subject to such a budget constraint. It is always easier and more pleasant to spend, as it were, “other people’s money” in an unaccountable (and hence inevitably somewhat irresponsible) fashion, which is no doubt one reason for the continued political attractiveness of the ability model of government.

⁷Rojas (1999) stresses the dominance of “controlled, monitored” decentralization in Latin America. Of

local services directly, or earmark and restrict local financing to such an extent that they might as well do so (Annex 2). In such circumstances, local citizens look to the central government to fix potholes, and they are right to do so. Just as children who are never given responsibility seldom act like responsible adults, this type of micro control generally produces weak and incapable local governments.

While less common, the opposite ill of under control has also emerged in some countries as a result of inappropriate decentralization strategies. For example, a number of East European countries in transition have given local governments large shares in national revenues as well as responsibility for important public service functions, without at the same time setting up an adequate institutional structure to ensure that central funds are being properly spent in, say, maintaining minimum standards of service in such areas of overriding national concern as education or health (Bird, Ebel, and Wallich, 1995). Similar concerns have frequently been expressed in some Latin American countries.

Finally, local governments frequently receive perverse signals from central governments. If, for example, the amount of transfers a government receives depends upon the size of its budget deficit, incentives are clearly perverse. As another example, national funding is often available for infrastructure investment but no funds are available for operation and maintenance (Annex 7). This creates a perverse preference for new investment and rewards localities for not maintaining existing facilities (which they would have to do out of their own funds) in order to strengthen their apparent need for new ones (for which the central government will pay).

Even when local government face perverse incentives, local efforts and policies may sometimes make a real difference. In Brazil, for example,

some cities are well-run and provide efficient services. Others, superficially similar in character and resources, are badly-run and poorly equipped. In Colombia, some departments provide superior health services compared with others whose resources are similar. Almost everywhere, some local governments do much better than others. The reason may simply relate to historical circumstance. For one reason or another an area started to do something well some time ago, and it continues to do so. Or it may be because of a caring and charismatic local leader or some other chance circumstance. Whatever the cause, such experiences highlight two important points. First, even in the perverse situation in which many local governments are placed by inappropriate central policies, there is usually some scope for local initiative. Second, such local initiatives may make a real difference in the lives of local people. One way to make decentralization effective and efficient is to make it easier for good examples to occur by giving subnational governments some room to maneuver, and, in part through facilitating an adequate flow of information, to make it easier to emulate their success.⁸

Many Latin American countries fail to recognize explicitly the diversity of local governments. "Local government" is a term that covers a wide range of realities. Cities of 10 million people, villages with 200 inhabitants, densely populated rural and urban areas, sparsely populated territories—all are, as a rule, organized in one form of local government or another. Some localities, are rich, some are poor. Some, have strong local community spirit, some have none. Some are run by well-intentioned, well-trained people; others by incompetent and corrupt officials. Unfortunately, the diversity of local government reality is seldom matched by equal diversity in central government rules. Even though some areas may manage to break out of the inappropriate mold into which they have been put, most will not. Successful decentralization must recognize the diversity and

course, as Carlsen (1998) and Schweger (1999) analyze rigorously, and as is discussed later in the present paper, there are often good theoretical and practical arguments for maintaining some degree of central control.

⁸As Rojas (1999) notes, dissemination of best practices and horizontal exchanges among subnational governments are probably more effective in improving local capacity than formal technical assistance: see also the discussion in Section 7.

heterogeneity of local government and allow for an equally diverse and heterogeneous set of responses to particular decentralization initiatives, including accommodating home-brewed solutions to particular local problems.

A useful approach in some circumstances may be through contracts or the making of specific agreements with different areas in accordance with their capacities and interests. The central government may, for example, contract with a particular local government (preferably for a period of years, and preferably in an open and mutu-

ally agreed fashion). To provide a particular service, as has to some extent been done in Colombia (World Bank, 1996a). At its best, this approach focuses pragmatically on what may work, rather than on trying to force everyone into the same centrally-determined Annex. Given the reality of diversity and the apparent political need in many countries to have nominally uniform laws, perhaps only something like a contract approach may be able to provide the necessarily non-uniform conditions needed to secure more or less uniform outcomes at least cost.

The Big Questions

Four big questions must be answered with respect to intergovernmental finance in any country:

(1) Who does what? —The question of *expenditure assignment*.

(2) Who levies what taxes? —The question of *revenue assignment*.

(3) How is the (virtually inevitable) imbalance between the revenues and expenditures of subnational governments that result from the answers to the first two questions to be resolved? —the question of *vertical imbalance*.

(4) To what extent should fiscal institutions attempt to adjust for the differences in needs and capacities between different governmental units at the same level of government? —The question of *horizontal imbalance*, or *equalization*.

Ideally, to achieve the relevant policy objectives, these questions should be approached within each country's specific circumstances. These objectives include not only the normal public finance trio of efficiency (allocation), equity (distribution), and stabilization but also economic growth as well as such nebulous but politically resonant goals as "regional balance" and maintaining national integrity and political stability. There may be conflicts between these objectives, as well as differences between local and central perceptions of the weights to be attached to them. Moreover, like all public policies, intergovernmental fiscal policies must take into account both political constraints (such as the strength of different regions and groups in political decisions) and economic constraints (such as the development of financial markets). Finally, all policy change must start from the here and now. Since every country has its own history, the current state of fiscal institutions in large part reflects the results of an accretionary process of policy change over time. Therefore, to understand, let alone to resolve, the intergovernmental fiscal puzzle in any country

thus requires substantial institutional as well as analytical knowledge.

Expenditure Assignment

How government functions should be divided among levels of government is a complicated matter, to which each country has its own unique solution. Section 4 sets out a few principles and alternative approaches with respect to the spending side of public sector budgets. The main point made is that the most important question about expenditures, from an economic perspective, is not *who* does what but rather *what* is done, and how well it is done. The details of expenditure assignment are less important in determining outcomes than the clarity of that assignment, how it is determined and implemented, and how well expenditures are managed at all levels of government.

Revenue Assignment

Who should levy what taxes, and how effectively they can do so, has been a major issue in some countries. The correct revenue assignment in a multi-level government structure is by no means clear in principle, and is likely to prove highly controversial in practice. The fundamental problems are two. First, the central government can inherently collect most taxes more efficiently than subnational governments. Second, the potential tax bases available to the latter vary widely from jurisdiction to jurisdiction. The first of these problems gives rise to vertical imbalance; the second produces horizontal imbalance.

The first of these problems could be solved to the extent that variable surcharges on central taxes are feasible. Given the paucity of such surcharges in Latin America, however, vertical imbalance may indeed give rise to concern. In addition, since the second problem, horizontal imbalance, would actually be worsened by more decentralized revenues, even if vertical imbalances were resolved by

adjusting revenue assignments, there would still be a need for transfers, at least in countries in which such interjurisdictional disparities are of policy concern.⁹

Closing the Gap

No matter how revenues and expenditures are assigned, worldwide experience suggests that some vertical imbalance is likely to remain. Even if the tax base of the richest subnational government enabled it to balance “own” revenues and expenditures, imbalances would remain for all the rest. Even for the richest region, the differential elasticity of expenditures and revenues assigned to different government levels may, in any case, soon lead to a re-emergence of a vertical imbalance problem. One way or another, structural “gaps” are likely to exist in all multi-tier fiscal systems.

In principle, there are at least four ways in which any subnational revenue-expenditure gap might be closed.

Revenues could be increased at the subnational level. In many Latin American countries, such increases seem both feasible and desirable (Bird, 1999b). But even if subnational revenues are increased as much as is politically and economically feasible, it may not be enough in some countries and will certainly not provide sufficient resources to finance adequate service levels in many poor localities.

Subnational expenditures could be reduced. While always popular with central governments and no doubt sometimes necessary, this approach is unlikely to be advisable if the system has been

properly designed and the right expenditures assigned to the subnational level in the first place.

Some expenditure functions could be transferred up to the level with more revenue (or revenue-raising power) or some revenues transferred down to the level with more expenditure. While such shifts would be both unnecessary and unwise if the basic structure of the system is correct, this is a big “if” in many countries.

Some centrally-collected revenues could be transferred to subnational governments. In the end, in virtually every country in the world, even those in which subnational governments have much greater access to revenues than in Latin America, there invariably remains some need for intergovernmental transfers. As discussed in Section 6, however, the design and implementation of transfers involves many more considerations than simply closing the subnational revenue-expenditure gap.

In practice, intergovernmental fiscal transfers are the major source of subnational revenue in most Latin American countries (IDB, 1997). This situation seems unlikely to change significantly in the near future. Together with a properly designed subnational tax base, a properly designed transfer system is, therefore, essential to ensure that adequate incentives for fiscal responsibility are provided. Unless made accountable in this sense, local decisionmakers are unlikely to make effective and efficient resource allocation decisions (to the extent that *any* political process can produce such decisions). In this respect accountability *at the margin*, and not simply the size of the amount transferred is critical. It is possible in principle for a provincial government to be almost totally dependent on central transfers and still be fully accountable, provided the amounts are fixed in advance and cannot be altered as a result of any (in-period) action by the recipient. On the other hand, a badly designed transfers, even if quantitatively much smaller, can do a lot of damage to the efficient and effective operation of the local public sector. A well-designed transfer can ensure that at the margin, local actions to raise or lower local revenues or expenditures will directly affect outcomes, which is what is needed to ensure political

⁹As May (1969) argues, the “taste” for regional equalization may vary greatly from country to country. Compare, for example, the explicit interregional redistribution of the unemployment insurance system in Canada with the state-based (and hence non-equalizing) system of the United States. Note that regionally redistributive transfers must be distinguished from so-called “equalization” transfers discussed in Section 6 (and also found in Canada, see Annex 5) which are intended to ensure specified efficiency objectives.

accountability.¹⁰ One should not underestimate, however, either the difficulty of implementing such a system in the real world of politics or the likelihood that the transfers currently in place in most countries fall far short of ideal.¹¹

Other Issues

Three additional issues have attracted considerable attention in recent discussions of fiscal decentralization in Latin America. The first relates to local capacity. (Are local governments up to the job?) The second focuses on the need or desirability of specific restrictions on subnational borrowing. And the third relates to the broader issue of the impact of decentralization on macroeconomic stability and policy. Important as these issues may be in particular context, in the long run they are secondary to the key questions of who

does what and how is it financed. If subnational tax policy and transfers are appropriately designed, the resulting hard budget constraints should, over time, curb any tendency to either microeconomic inefficiency or macroeconomic instability, hence largely vitiating the need for special concern about subnational borrowing. Given, however, that reality in many countries is often far from ideal and that there may be substantial political difficulties in altering it (Salegh and Tommasi, 1999), “second-best” policies may be needed areas. The capacity issue too can and should be addressed best by a combination of good initial design and the establishment of an appropriate institutional framework, even though political interests may block such moves in the near future in many countries.¹² These points are discussed further in Sections 7 through 9.

¹⁰As discussed in Section 6, two special forms of transfers also need to be considered. First, some transfers are explicitly intended either to ensure the provision of specific services at specific levels (as in the case of the capitation grants discussed in Annex 6, in which the transfers are essentially payments to agents). Second, transfers may be explicitly intended to induce subnational governments to provide more of certain services than they would otherwise do (as in the discussion of matching grants for infrastructure in Annex 7).

¹¹See, for example, the detailed analyses of transfers in Colombia and Argentina in World Bank (1996a, 1996b).

¹²For example, Rojas (1999) notes that no country in Latin America has thus far managed to create adequate technical unit to monitor and foster decentralization: see Section 9 for further discussion.

Government Spending

The basic requirement for efficient and effective subnational government may be thought of as the “matching principle.” As already discussed, to the extent possible, benefit areas should be matched with financing areas as in the benefit model of local finance. In addition, expenditure responsibilities should be matched with revenue resources, and, finally, revenue capacities should be matched with political accountability.

Assigning Expenditures

The basic rule of efficient expenditure assignment is often taken to be to assign each function to the lowest level of government consistent with its efficient performance.¹³ So long as there are local variations in tastes and costs, there are efficiency gains from carrying out public sector activities in as decentralized a fashion as possible.¹⁴ Indeed, from this perspective, the only services that should be provided centrally are those for which there are no differences in demands in different localities, where there are substantial spillovers between jurisdictions that cannot be handled in some other way (by contracting, by redrawing boundaries, or by grant design), or those for which the additional costs of local administration are sufficiently higher to outweigh its advantages.

Under this approach, apart from the important issue of distribution, almost all public services (other than national defence, foreign policy, and surprisingly few others) should in principle be delivered at the local level. Decisions about the services provided, to whom they are provided, and in what quantity and quality, should be made locally and local taxpayers should pay for the services provided. In practice, however, although some functions (such as street maintenance) are

local everywhere, the allocation of functions to subnational governments varies considerably from country to country (IDB, 1997).

This approach to expenditure assignment need not be consistent with the matching principle, since there is nothing that guarantees that the bundle of services thus assigned to any particular level of government will be matched by the set of revenue instruments assigned to that same level. On the contrary, as already noted in Section 3, a fundamental imbalance emerges in the vertical assignment of expenditures and revenues in virtually every country. Hence, there is an apparent need for intergovernmental fiscal transfers to close the budgetary gap.

As emphasized in Section 2, the essential economic role of local government is to provide to local residents those public services for which they are willing to pay. Local governments are (or at least should be) accountable to their citizens to the extent that those citizens finance those actions. Such accountability is the public sector equivalent of the “bottom line” in the private sector. Three conditions need to be satisfied to achieve accountability in this sense in subnational finance. First, subnational governments should, whenever possible, charge for the services that they provide (Annex 3). Second, if charging is impracticable, subnational governments should finance such services from taxes borne by local residents (Section 5), except to the extent that the central government is willing to pay for them (Section 6). Third, where the central government does pay as a rule, subnational governments should be accountable to the central government (to at least some extent) when the central government pays (Section 9).

¹³In the context of the European Union, this has come to be called the “subsidiarity principle:” see Breton, Cassone and Frascini (1998) for a useful discussion.

¹⁴Oates (1972) has labelled this “the decentralization theorem.”

To apply these rules, an essential first step is to establish clear lines of responsibility and accountability.¹⁵ But *clarity* of assignment in terms of specifying exactly what services, each governmental agency is responsible for, is only part of the story (Annex 1). Clarity must be matched by *accountability*, in terms of political democracy and transparency of operation, as well as by *authority* in terms of both the ability to manage expenditures and to determine (within limits) revenues. Moreover, even in the best of all possible worlds full clarity in expenditure assignment may never be fully attainable, for several reasons.

First, with respect to many important spheres of public sector activity (for example, education) different jurisdictional levels may play critical roles. The central government may, for example, appropriately set national standards for graduates and for teachers and may also establish the basic curriculum to be covered. Within this framework, regional governments may develop their own policy goals (for instance, with respect to school facilities) and deploy appropriate regulatory instruments to achieve them. Local governments may be responsible for actually paying teachers and maintaining facilities. Finally, educational services are delivered by local schools which may often, experience suggests, produce better outcomes if they have a substantial degree of budgetary autonomy and can react to input from teachers, parents, and the local community (Burki et al., 1999). Three or more levels of government may thus play important roles in delivering educational services.

It is not meaningful to consider a particular service as assigned to the local level, when much of the relevant policy and regulatory framework, and indeed much of the financing, comes from higher levels of government (and delivery of the service may take place at a lower institutional level). Clarity as to whom is responsible for what is indeed important, but many different public sector actors may appropriately be involved in the provision of any particular public service. What matters is not so much that each expenditure function,

¹⁵This is, for example, one of the key principles set out in the IMF's recent Code of Good Fiscal Practice.

broadly understood, is clearly assigned to one level of government or another, but rather that it is clear to all *exactly* who is responsible for doing precisely what.

Second, clarity in expenditure responsibility is sometimes taken to mean that much attention needs to be paid to the problem of coordination. In this regard, however, it should be recalled that the principal argument *for* decentralization is that coordination (or cartelization, or monopoly, as it might also be called) is not delivering the goods, or at least not delivering the right goods in the right quantities to the right people. Moreover, what may at first glance appear to be undesirable duplication or overlapping of functions may, in some instances, reflect either useful redundancy in a complex system (Landau, 1969) or desirable governmental competition (Breton, 1996). On the other hand, in the real world duplication and confusion may often imply waste and the need for better intergovernmental coordination. The answer, however, is not to abandon decentralization but rather to design intergovernmental fiscal relations in such a way as to minimize real coordination problems and to continue working at the difficult task of establishing effective coordinating institutions (Annex 8).

However, even the best-decentralized public sector is undoubtedly far from a perfectly competitive market structure and outcomes will never be optimal in the technical economic sense.¹⁶ The most appropriate approach is to establish budget constraints that are as hard as possible for all relevant decisionmakers and to make the operation of the system as transparent as possible. "Letting 100 flowers bloom" in the form of relatively uncoordinated decentralized public sector suppliers striving to meet clearly specified and publicly accountable mandates may, in the end, provide a better laboratory for the development of new and better public sector services than any conceivable centralized alternative (Oates, 1998).

¹⁶Of course, as Goodin (1996) emphasizes, how important the failure to achieve economic optimality is depends upon the extent to which policymakers are actually trying to do so.

Managing Expenditures

No matter where subnational governments get their funds, they are unlikely to have enough to do all they or their citizens want and expect. Successful local government requires that scarce public funds be managed as efficiently and used as effectively as possible. Both financial honesty and political accountability require that budgeting and financial procedures are properly established and implemented. Budgeting, financial reporting, and auditing should be comprehensive, comprehensible, comparable, verifiable, and public.

It is equally important, however, to ensure that budgeted resources are applied as efficiently and effectively as possible to achieve desired public outcomes. Adequate and appropriate procedural norms are important in any financial system, but substantive outcomes are what really matters. Proper public expenditure management must adequately control the total level of revenue and expenditure, appropriately allocate public resources among sectors and programs, and ensure that governmental institutions operate as efficiently as possible (World Bank, 1998). It seems particularly critical in this respect for subnational governments to have sufficient *authority* to manage both expenditures and revenues sides.

For this reason, it is important to ensure that correct incentives exist on the revenue side. Specifically, effective subnational governments must have significant revenue sources under their control, for which they are economically and politically responsible. They must also be able to predict with considerable certainty anticipated intergovernmental transfers in any financial period.¹⁷

In many Latin American countries, decentralization has been recently characterized by the earmarking of substantial parts of intergovernmental transfers to localities to local infrastructure investment. This is the case in Argentina (for

housing), Brazil, Colombia, Chile, Ecuador, Guatemala, and Venezuela. Presumably motivated in part by the desire to prevent local governments from “wasting” transfers on expanding local payrolls, earmarking has had the paradoxical effect of exacerbating local fiscal problems in some instances (Annex 2). Although not fully effective (since money is fungible, there is usually some substitution of transfers for own-source revenues) the result has usually been to expand capital spending to some extent, while making the already difficult problem of funding the operation and maintenance of these investments even more difficult. Even when it succeeds in fostering local investment, earmarking has little to be said in its favor. It distorts local preferences, exacerbates perverse incentives already found in the local finance system, and often connects revenue sources with expenditures in totally illogical ways. Moreover, excessive earmarking (like the related process of “mandating” subnational governments to spend in accordance with central preferences rather than their own) significantly reduces the scope of subnational governments for effectively managing expenditures, even if they had both the will and the capacity to do so.

Good subnational budgeting should take place within the framework of medium-term expenditure framework (MTEF) both to ensure the proper financing of investment projects and to reduce the scope for the short-term political manipulation of budgets (for example, to expand pre-election public employment in an unsustainable fashion).¹⁸ An essential first step in this direction is to put sound budgetary and financial procedures into place, especially in the more important subnational governments such as states and large cities.

Subnational budgets, like central budgets, should be comprehensive, accurate, periodic, authoritative, timely, and transparent. The budget law

¹⁷It is thus not appropriate to have exactly the same budget period for local and central governments, since the former cannot determine their budgets until they know the size and nature of the transfers they are likely to receive from the latter.

¹⁸Unless there is an adequate MTEF at the *central* level, it may be difficult to require one locally. On the other hand, given the growing importance of local governments in many countries, it is increasingly difficult to have an adequate MTEF at the central level without explicitly incorporating aggregate local revenues and expenditures into the budgeting and planning exercise.

should be uniform and clear, and it should be enforced. Moreover, expenditures should be subject to external audits to ensure that the law is followed. All this would seem to require is a strong central hand to ensure compliance with the rules. For example, the central government should establish a “framework” budget law and require adequate external audits (such as by a private sector firm). It should not, however, subject subnational budgets to prior approval, or the whole point of decentralization is lost.

A strong budgeting and financial system along these lines will satisfy two essential requirements of good government. First, it will establish the basis for financial control. Second, it will provide reasonably accurate, uniform, and timely financial information. Even the best set of financial procedures does nothing, however, to ensure that scarce public resources, even if properly spent and accounted for according to law, have been spent in the best possible way or as efficiently as possible. Nor, as some countries have seen in recent years, does even the best-enforced set of budgetary procedures ensure that aggregate fiscal discipline will be adequately maintained. To attain favorable outcomes in these respects, additional important fiscal institutions need to be put in place.

For example, with respect to *fiscal discipline*, the MTEF mentioned above can provide an essential baseline for measuring and monitoring the budgetary impact of policy changes. Similarly, it is critical for fiscal discipline that the amount of revenue expected from intergovernmental transfers is predetermined and is not subject to political re-negotiation during the budget year. Moreover, subnational governments must not be able to depend on central government bailouts of imprudent financial decisions, such as unsustainable borrowing or expenditure increases. On the other hand, subnational governments should be able to increase expenditures to the extent that the full fiscal consequences of such increases are borne by local residents and, equally importantly, they should be able to *reduce* expenditures if they so wish.¹⁹

¹⁹As Rodden (1999) notes, Brazilian states are constitutionally restricted in the extent to which they can

Similar institutional safeguards should be in place to ensure allocative and operational efficiency within affordable fiscal limits. *Allocative efficiency* requires that managers at all levels receive adequate and accurate information on the effectiveness and social outcomes of the programs for which they are responsible. This can take place for example, through the revenues produced by properly-designed user charges (Annex 3) and/or through participatory interaction with clients at both the budgetary and implementation stages (Vergara, 1999). Moreover, they must have strong incentives to respond to these signals, for example, by facing a predetermined spending limit which can be altered only if they can “sell” more services that their client groups are willing to pay for.

Operational efficiency may best be achieved in most cases by allowing line managers (like the school managers mentioned earlier) significant discretion, within budgetary cash limits, to reallocate funds among inputs or perhaps even across budgetary periods, subject to the attainment of predefined operational (performance) goals as well as compliance with appropriate internal and external financial control and audit systems. Emphasis is thus shifted from input controls (hiring so many persons at such and such a wage or renting so many square meters of space) to output controls (providing health care of a determined quality to so many persons within a specified time period or issuing so many marriage licenses).

Shifting the emphasis in public finance from inputs to outputs in this way is an essential step in improving policy outcomes at any level of government. It carries with it some risks,²⁰ and it is by no means as yet fully clear how best, or to what extent, this shift can be accomplished in the difficult circumstances facing most countries of the region. Nonetheless, there is already considerable experience in Latin America with such

reduce expenditures, even those funded solely from their own revenues. Such limitations are clearly not conducive to good expenditure management.

²⁰For example, there must still be full accountability by clearly identifiable decisionmakers with respect to all expenditure decisions to reduce the possibility of fraud.

techniques as client surveys, participatory budgeting (notably in some Brazilian cities), performance budgeting, and user financing (Burki et al., 1999). The path to success is by no means clearly marked. But at last, it seems, we are beginning to

see more clearly some of the directions in which countries should begin to move in terms of expenditure management if their leaders are seriously interested in making life better for their populations.

The Revenue Side

Before discussing the appropriate design of subnational taxes, it may be useful to clarify the meaning of a subnational tax. A *completely* subnational tax might be defined as one that is assessed by subnational governments, at rates decided by subnational governments, which is also collected by subnational governments, with its proceeds accruing to subnational governments. In the real world, however, many taxes may possess only one or two of these characteristics, and the “ownership” of the levy may be unclear.

In Argentina, for example, although part of the proceeds of many taxes accrue to the provinces, the rates (and bases) of these taxes are determined by the national government, which also assesses and collects them. For most purposes, such taxes are best thought of as central government taxes that are allocated to the provinces through transfers. This interpretation is particularly plausible because there is little connection between the amount transferred and the amount collected locally.

Nevertheless, what appears as a central tax with a related transfer program may, from some perspectives, be a subnational tax. If, for example, a provincial government decides whether or not to impose a particular tax, determines the tax base, sets the tax rate, and receives all the revenues, then even if the tax is collected by the central government, the only role the latter plays is as a collection agent. Presumably, the central government has a comparative advantage in tax collection, and the subnational government has contracted for its services in this respect. In this case, there is no intergovernmental transfer at all, except in the narrowest accounting sense.²¹

²¹Argentina’s revenue-sharing system has sometimes been interpreted along these lines because the provinces have delegated much of their revenue-raising power to the center and are, so to speak, simply being “compensated” through the so-called *coparticipación* transfer. As already noted, however, this interpretation does not appear to be economically meaningful both

Intermediate cases between these extremes may easily be found. In Brazil, for example, the states impose and collect their own VATs (ICMS), but the rate of the tax is set centrally and uniformly. Most Canadian provinces do not levy personal income taxes, but impose surcharges on the federal income tax, which is collected by the federal government and remitted to the provinces. A third example is the case of Russia where although the VAT base and rates are set centrally, proceeds are shared with the regional governments on a derivation basis. In fact, the VAT is actually collected locally and the (variable) federal share is remitted (not always in a timely fashion) to Moscow. In which of these three cases is subnational tax power the strongest? Theory and experience suggests that it is in Canada, the only one of the three in which the subnational governments do not, in fact, collect the tax in question. Why? Because the most critical aspect of subnational taxing power is who is politically responsible for setting the tax *rate*.

Tax Assignment in Principle

The traditional theory of fiscal federalism prescribes a very limited tax base for subnational governments. The only good local taxes are said to be those that are easy to administer locally, that are imposed solely (or mainly) on local residents, and that does not raise problems of “harmonization” or “competition” between subnational governments or between subnational and national governments.²² The only major revenue source that passes these stringent tests is the residential property tax. Taxes on vehicles and user fees may also fall under this category (Annex 3). Since this

because the revenues are not distributed on a derivation basis and (especially) because the relevant tax rates and bases are determined by the national government.

²²The classic tax assignment arguments are set out in Musgrave (1983). Recent restatements may be found in Oates (1998) and McLure (1999a). For a review of this literature, see Bird (1999a).

conclusion accords with the reluctance of most central governments to provide subnational governments access to their lucrative sales and income tax bases, it is not surprising that it has become conventional wisdom. Subnational governments almost everywhere are urged to make more use of property taxes, and criticized when they do not do so enthusiastically.

Up to a point, there is much to be said for this argument. Unfortunately, that point falls far short of the task facing many subnational governments for a number of reasons. First, the conventional case for property taxes is to some extent flawed. Such taxes are costly and difficult to administer well, and these problems are greatly exacerbated as the tax burden increases. Moreover, in practice, political reality means that increases in property taxes are often concentrated primarily on those nonresidential properties which most lend themselves to tax exporting, thus undercutting one of the principal arguments for local use of this tax base.

Second, even a well-administered local property tax cannot finance major social expenditures (education, health, social assistance) except perhaps in the richest (and usually largest) communities. To the extent that it is desirable for governments to finance from their own revenues the services they provide, either local governments financed by property taxes are confined to providing such local services as street cleaning and refuse removal, or they are heavily dependent on transfers from higher levels of government. This is the pattern in most developed countries, including the relatively few in which the property tax is the mainstay of local finance. In the OECD countries, to the extent that local governments are not dependent on national transfers, they invariably impose significant direct taxes on businesses or levy surcharges on national income taxes (Bird and Slack, 1991).

Third, the conventional argument does not take adequately into account the existence, in most larger countries, of important regional (intermediate) levels of government which often play a major role in financing social expenditures. Even if *local* governments, can to a considerable extent,

finance local services through property taxes and user fees on residents, *regional* governments that are responsible for such social services as a rule cannot. The conventional approach, while not firmly grounded in anything except expediency, has been to suggest that such governments are, in principle, best financed by retail sales taxes. Such taxes are usually assumed to fall mainly on residents.²³ As the experience of the United States and Canada have long demonstrated, retail sales taxes can be administered at the regional level. However, it seems unlikely that this experience can be applied to developing countries, which have universally found it impossible to administer such taxes even at the national level.

Given the recent move towards decentralization in Latin America and the concern frequently expressed about the resulting strain on intergovernmental fiscal relations and the possibility of irresponsible behavior by some subnational governments (Tanzi, 1996), the issue of subnational revenues requires re-examination. International experience suggests that subnational governments carrying out important expenditure functions are more likely to do so responsibly the more responsible they are for raising the revenues they spend. While there will obviously always remain an important role for intergovernmental transfers, especially in countries with wide regional economic disparities, there seems to be no reason why wealthier regions (including metropolitan areas) should not be able to raise and spend most of their budgets themselves.²⁴

²³In reality, most retail sales taxes fall to a considerable extent on business inputs even in developed countries. One Canadian study, for example, found that between one-third to one-half of the retail sales tax base in different provinces consisted of such inputs (Kuo, McGirr, and Poddar, 1988).

²⁴For an earlier statement of this argument, see Bird (1993). Note that an important implication of strengthening subnational revenues is that the resources accruing to different states or provinces will differ greatly, depending upon their access to the tax base in question. While transfer systems could, in principle, be adjusted as desired to prevent unduly penalizing poorer regions, as noted earlier, the extent to which such adjustments will actually be made is country-specific.

Two basic principles of assigning revenues to subnational governments may be suggested. First, “own-source” revenues should ideally be sufficient to enable at least the richest subnational governments to finance from their own resources all locally-provided services primarily benefitting local residents. Second, to the extent possible, subnational revenues should be collected only from local residents, preferably in relation to the perceived benefits they receive from local services. This approach is in turn based on three simple principles: First, more attention should be paid to matching expenditure and revenue needs at different levels of government. Second, greater effort should be made to ensure that all governments bear significant responsibility at the margin for financing the expenditures for which they are politically responsible. And third, subnational taxes should not unduly distort the allocation of resources.

The characteristics that might be sought in a subnational tax satisfying these requirements may include the following:

- The tax base should be relatively immobile, to allow local authorities some leeway in varying rates without losing most of their tax base.
- The tax yield should be adequate to meet local needs and sufficiently buoyant over time (that is, it should expand at least as fast as expenditures).
- The tax yield should be relatively stable and predictable over time.
- It should not be possible to export much, if any, of the tax burden to nonresidents.
- The tax base should be visible to ensure accountability.
- Taxpayers should perceive the tax to be reasonably fair.
- The tax should be relatively easy to administer efficiently and effectively.

Not everyone would agree that all these characteristics are necessarily or equally desirable. For example, is it unequivocally good that subnational governments should be insulated from either the tax base consequences of their tax rate choices or from inflation? More importantly, the characteristics that may be sought in an ideal local tax from the point of view of local and central governments are not necessarily compatible. Both levels might agree that the tax base should be immobile, and perhaps also that the tax yield should be stable and adequate to meet local needs. Yet, while central governments should be concerned about ensuring that the tax burden cannot be exported to nonresidents and that the local tax base is visible to ensure accountability, such attributes are less likely to be appealing to local governments.

The bottom line, however, is that unless local governments have some significant degree of freedom to alter the level and composition of their revenues, neither local autonomy nor local accountability are meaningful concepts. In particular, as noted earlier, subnational governments should be able to set tax *rates* (albeit perhaps within limits). Though this condition is seldom satisfied in developing countries, such rate flexibility is essential if a tax is to be adequately responsive to local needs and decisions, while remaining politically accountable.

Local governments should not only have access to those revenue sources that they are best equipped to exploit (such as residential property taxes and user charges for local services) but they should be both encouraged and permitted to exploit these sources. A potential danger in permitting local governments’ even limited freedom to tax is that they will not utilize fully all the revenue sources open to them, thus allowing the level and quality of public services to deteriorate. If intergovernmental fiscal structures are properly designed, this should not be a real problem. If the service in question is one of national importance (such as research, for example) or one in which there is a strong national interest in maintaining standards (for example, poverty alleviation), it should presumably be funded and monitored by the central government. If it is not a matter of national inter-

est, why should the central government be concerned? Of course, this argument assumes that the central government has both adequate information and a predetermined transfer system, as discussed further in Sections 6 and 9 below.

If local electors do not like what their local government does, or does not do, they can (try to) “throw the rascals out” at the next election. The freedom to make mistakes, and to bear the consequences of one's mistakes, is an important component of local autonomy in any country. Indeed, unless local governments are given some degree of freedom with respect to local revenues, including the freedom to make mistakes (for which they are accountable to voters), the development of responsible and responsive local government is likely to remain unattainable. Of course, if the conditions of effective democracy and adequate information are not satisfied, or if those who fail to collect local taxes, or to spend revenues efficiently are bailed out by discretionary transfers, the “rascals” may not be thrown out but rather re-elected for their success in obtaining a larger share of other people's money. Countries that, for whatever reason, fail to set up an appropriate intergovernmental fiscal structure are likely to have both more problems in managing decentralization and less satisfactory policy outcomes.

Another danger is that local governments may attempt to extract revenues from sources, for which they are not accountable, thus obviating the basic efficiency argument for their existence. To avoid this problem, it may be desirable to limit local government access to taxes that fall mainly on nonresidents (such as most natural resource levies, pre-retail stage sales taxes and, to some extent, nonresidential real property taxes). One way to deal with this problem may be to establish a uniform set of tax bases for local governments (perhaps different for different categories such as big cities, small towns, and rural areas), with a limited amount of rate flexibility permitted in order to provide room for local effort while restraining unproductive competition and unwarranted exploitation. If inappropriate tax bases are assigned to subnational governments, wasteful competition and undesirable tax exporting are likely to result.

These general points are discussed further with respect to two distinct, and important problems in subnational taxation. The first is how to finance regional governments. This problem is especially important in larger and federal countries, but it is by no means confined to them. The second problem, found everywhere, is how to finance “local services” more narrowly understood.

Financing Regional Governments

The present assignment of taxes in most countries with important regional levels of government, such as Brazil and Argentina, seems particularly deficient. One problem is that there is a significant vertical imbalance between expenditures and revenues, with consequent implications for autonomy, efficiency, and accountability. Another problem is that the present confused and confusing system results in significant costs—costs of administration, costs of compliance, and costs arising from tax-induced inefficiencies in the allocation of scarce resources.²⁵

In principle, multitiered governments work best when taxes and the benefits of public spending are as closely related as possible—when citizen-voter-consumers residing in a particular political jurisdiction pay for what they get from the public sector and get what they pay for (that is, benefit from the expenditures financed by the taxes they pay). Obviously, when citizens reside in several overlapping jurisdictions (local-state-nation) this so-called “principle of fiscal equivalence” (Olson, 1969) suggests that they should pay taxes to each

²⁵For a detailed discussion of the latter problem, see Bird and Mintz (1999). A particular problem arises in some countries because of the uneven geographical distribution of natural resources and the resulting severance of the link between local taxes and benefits when subnational governments are able to tax such resources, as is often the case. The ideal solution is to prevent them from doing so (Mieszkowski, 1983), but if this is not possible, considerable care must be taken in designing other aspects of intergovernmental finance, particularly transfer systems, in order to offset the resulting distortion as much as possible. Unfortunately, although this problem is important in a number of countries in the region, it cannot be discussed further here.

level corresponding to the benefits they receive from each jurisdiction. In this framework, inter-governmental transfers would exist to restore this equivalence when, for example, some benefits flow from one jurisdiction to another or (negatively) when some taxes levied by one jurisdiction are paid by persons residing in another.²⁶

Considerations of administrative efficiency and feasibility may dictate that higher (or lower) levels of government impose certain taxes or carry out certain expenditures even when it would not be strictly appropriate to do so on equivalence grounds. Vertical fiscal flows, like those that dominate the intergovernmental fiscal scene in most countries, are motivated largely by this consideration, at least with respect to those flowing to richer jurisdictions. In contrast, if more adequate subnational taxes are made available, this “fiscal gap” (Boadway and Hobson, 1993) argument for transfers disappears. There is then no case for universal intergovernmental fiscal transfers, since in this system the richest units of government at subnational levels should be essentially self-sufficient (Bird, 1993). Any grants from higher levels of government made for reasons of regional equalization in this system should then be clearly *inframarginal* so that, as McLure (1999a) notes, all subnational governments face the full marginal tax price of the spending decisions for which they are responsible, thus yielding the hard budget constraint emphasized by such authors as Tanzi (1996).

Good subnational taxes (at both regional and local levels) should thus satisfy two main criteria. First, they should provide sufficient revenue for the richest subnational units to be essentially fiscally autonomous.²⁷ Second, they should clearly impose fiscal responsibility at the margin on subna-

tional governments. As mentioned earlier, the simplest and probably best way to achieve the latter goal is by allowing them to establish their own tax rates with respect to, at least, some major taxes.

The most immediately important subnational revenue issue facing large countries in Latin America is the development of a satisfactory revenue base for regional governments, that is, one for which they are politically responsible. While more can be done in the form of regional excise taxes, especially on vehicles and fuels, in most countries (Bird, 1999b), if regional governments have significant expenditure responsibilities, there are really only two important possibilities: a surcharge on the central personal income tax (PIT) or a surcharge on the central value-added tax (VAT). If local governments are to be *both* large spenders *and* less dependent on grants, they must have access to national tax bases. “Piggybacking” through surcharges seems to be the only viable way to do this while retaining an important element of political accountability.

The possibility of local income tax surcharges now seems to be broadly accepted. Unfortunately, as Shome (1999) has recently emphasized, few Latin American countries have sufficiently robust central income taxes to offer much hope that subnational governments will soon be able to derive much revenue from this source. A potentially more promising alternative for subnational revenues may thus turn out to be a surcharge on the VAT. Such a tax already exists and works well in Canada, and its implementation now seems feasible even in countries with less well-developed tax administrations (Annex 4). At least in the larger Latin American countries with federal features, such as Brazil, Argentina, or Mexico, this path to increased regional taxation seems to warrant further exploration.

Financing Local Services

Turning to local taxes, apart from user charges (Annex 3), there seem to be only two major possibilities—a revised, and revived, property tax and an improved form of local business taxation.

²⁶Note that such transfers would be horizontal, between provinces or municipalities, and not between levels of government. An example of such a system (at least in principle—the practice is quite deficient from this perspective) is the *Fondo Común* in Chile.

²⁷This does not preclude intergovernmental fiscal transfers to achieve the usual “spillover” objectives but also, as noted in Section 6, to ensure the adequate provision of certain services to “national standards”

The Property Tax. Almost without exception, much can and should be done to strengthen the deficient property taxes now in place in most Latin American countries. The tax should be simplified and applied uniformly. Cadastral maps should be updated and valuations made more consistently and currently. Improved use should be made of flows of information from property registries, local building license authorities, public utilities, etc. In addition, as Dillinger (1991) has argued, from a revenue perspective, more attention should be paid to improving the “sharp end” (collection and enforcement) rather than to the technically more costly (and less immediately productive in terms of revenue) mapping and surveying of the traditional cadastral approach.

Nonetheless, property taxes are not easy to administer, particularly in countries where inflation is endemic, and they are never politically popular owing to their visibility and to certain inherent administrative difficulties. Even in the most sophisticated countries, local property taxes seldom yield enough to finance local services. No developed country which depends significantly upon property taxes for local fiscal resources have a local government sector that accounts for more than 10 percent of total public spending (Bird and Slack, 1991). Similarly, property taxes seldom account for more than 20 percent of local current revenues (or less than 1 percent of total public spending) in developing countries (UNCHS, 1996). The property tax is a useful, even necessary, source of local revenue, but it cannot easily provide sufficient resources to finance a significant expansion of local public services in most countries. Indeed, many countries have been hard-pressed even to maintain the present low relative importance of property tax revenues in the face of varying price levels and political difficulties.

These facts reflect both the political and the administrative realities of property taxation. One reason for the commonly observed phenomenon of widespread resistance to the property tax, for example, is simply its visibility. The tax generally has to be paid directly by taxpayers in periodic lump sum payments. Taxpayers who pay taxes directly to the government tend to be more aware

of the size of their tax bill than those whose take-home pay is reduced by weekly or monthly tax deductions. The need to make large periodic payments may well add to the accountability and responsibility of governments, but it also increases the sensitivity of taxpayers to even nominal increases in taxes.

The inelasticity of the property tax has a similar effect. Since the base of this tax does not increase automatically over time, the periodic nominal increases in property tax bills needed to maintain real revenues when price levels rise require increased tax rates. In terms of political accountability, the need to confront the people with the cost of government is a virtue of the property tax. However, the downside, at least from the government's point of view, is the heightened visibility of nominal tax increases and the accompanying political resistance.

Finally, property taxes best finance such services as roads and garbage collection. The quantity and quality of these services (or their absence) is thus readily linked to the property tax. When potholes develop in their street, taxpayers are understandably quick to question the taxes that supposedly finance street repair. Again, the very feature that makes the property tax a good source of local government revenue also makes it especially vulnerable to political resistance.

Other problems result from property tax administration. As a rule, for example, property is assessed on the basis of its market value, usually defined as the price struck between a willing buyer and a willing seller in an arm's length transaction. Even in countries with well-developed property tax systems, political and technical issues lead to discrepancies between assessed values and market values within classes of property, between classes of property, and across municipalities (Bird and Slack, 1993). Since taxpayers can easily compare their property taxes with those of similar properties in their neighborhood, such discrepancies lead both to specific assessment appeals and to general pressure for tax relief.

For such reasons, experience around the world suggests that the political cost of reliance on resi-

dential property taxes is so high that no government with access to “cheaper” sources of finance will willingly do so. Both intergovernmental transfers, which can be spent as local governments wish, and local taxes on business, which can largely be exported, must, as a rule, be curtailed if local citizens are to be confronted more openly (through their local tax bills) with the true economic (and political) costs of the decisions made by those whom they have elected.

In particular, the temptation to indulge in politically painless but economically inefficient “tax exporting” suggests that constraints should be placed on the local taxation of nonresidential property.²⁸ As discussed below, some local business taxation may be justified on benefit (efficiency) grounds, but it should always be strictly constrained in order to preclude localities from attempting to shift the costs of services to outsiders.

Other policy reforms are needed to turn the property tax into a responsive instrument of local fiscal policy. First and most importantly, as emphasized earlier, local governments must be allowed to set their own tax rates. Few countries currently give local governments much freedom in this respect. Secondly, the tax base must be maintained adequately. In countries with much inflation, some form of index adjustment may be advisable (World Bank, 1989). National assessing agencies must, if necessary, be provided direct financial incentives to keep the tax base up to date. Finally, procedural reforms are often needed to improve collection efficiency, valuation accuracy, and the coverage of the potential tax base. None of these steps is easy, but countries that want to have local governments that are both responsive and responsible must follow this difficult road. There are no short cuts to successful local property taxation.

²⁸There appear to be no studies of this phenomenon in Latin America. Studies in the U.S. (McLure, 1967) and Canada (Ballentine and Thirsk, 1982), where residential property taxation is much more important than in Latin America, suggest that “exporting” local taxes on business is both a common and an important phenomenon. It seems unlikely to be less so in Latin America.

Local Business Taxes. Another critical problem in many countries is the reform of various unsatisfactory subnational taxes on business. While the ability to distort market conditions through such taxes must be restrained (for example, by establishing a uniform national base for local business taxation, with a minimum and maximum rate²⁹) there is both an economic (benefit) case for some regional and local taxation of business and, it seems, often an overwhelming political need for local leaders to impose such taxes. Indeed, in many countries, given the restrictions on (and unpopularity of) residential property taxes and the unreliability of central transfers, local business taxes may be the only way in which local governments are able to expand revenues in response to perceived local needs.³⁰

Unfortunately, most forms of local and regional business taxes found in developing countries may introduce serious economic distortions. These include corporate income taxes, trade taxes, business taxes, differentially heavy nonresidential property taxes, and even so-called “retail” sales taxes (usually levied in the form of levies based on estimated gross receipts). One way to reduce such problems is through a so-called business value tax (BVT)—a relatively low rate flat tax levied on an income-type value-added base (Annex 4).³¹ In contrast to the earlier suggestion of a regional VAT surcharge, which was motivated mainly by the desire to provide more adequate revenues to regional governments and hence to encourage greater fiscal responsibility and accountability, a BVT is aimed primarily at

²⁹A minimum rate is needed to prevent distorting “tax competition” (with richer jurisdictions, those with larger tax bases, lowering rates to attract an even larger tax base). A maximum rate is needed to prevent distorting “tax exporting” (as when jurisdictions in which breweries or gas distribution pipelines are located impose especially heavy taxes on such facilities in the expectation that the taxes will ultimately be paid by persons not resident in the jurisdiction).

³⁰Another, equally undesirable, outlet has sometimes been borrowing, as discussed further in Section 8.

³¹The history of this idea, and various partial examples found around the world, are set out in detail in Bird (1999a). An empirical application of this approach to Canada may be found in Bird and Mintz (1999).

improving the allocative efficiency of subnational revenue systems. Therefore, it seems less likely to find a welcoming political audience. Nonetheless, this line of action may offer a potentially promising alternative to the proliferation of increasing, and distorting, subnational business taxes that otherwise seem likely to lie in the future of many countries.

The Future of Tax Assignment

Suppose that both of the new proposals for subnational revenue reform mentioned in this section were to be accepted in any given country. The result would then be a “family” of VATs, with a standard invoice-credit destination-principle consumption-type VAT imposed at the central government level, corresponding state VATs imposed at varying rates on the same base by regional governments, and a BVT (essentially an income-type VAT imposed by the subtraction or the addition method) levied on all VAT registrants by regional (and perhaps larger local governments) at relatively uniform rates. In addition, all levels of government should apply appropriate user charges and some excise taxes (particularly those related to vehicles) might appropriately be levied at the regional level. There is also a strong case in most developing countries for more effective local taxation of residential property. Finally, the central government may continue to levy both a corporate and a (progressive) personal income tax, perhaps with regional governments imposing flat-rate PITs on the same base.

Much work needs to be done to develop the details of such a tax assignment scheme in specific national circumstances. Some of the issues that need consideration are the myriad details of design and administration, the relation between different levels of subnational government and the role and design of intergovernmental fiscal transfers.³² The appropriate and tolerable level of

³²Devolving more taxing authority to subnational governments will mean most in revenue terms to those with the largest tax bases. As noted earlier, if a country is at all concerned with providing relatively equal opportunities for access to key public services, a more explicitly equalizing transfer system may be required.

asymmetry in the application of the suggested principles to subnational governments of vastly differing size and competence also needs to be established. Nonetheless, it does appear that the three long-accepted principles governing subnational taxation detailed below need to be discarded.

First, the conventional model of tax assignment, which in effect assigns all significant revenue sources to central governments, is clearly inappropriate for countries in which subnational governments, for whatever reason, account for a significant proportion of public sector spending. If such governments are to be big spenders, they must, in the interests of fiscal responsibility and accountability, also levy higher taxes.³³

Second, the VAT is the key to central government finance in most developing countries. Central governments are most reluctant to lose control over this tax, and this understandable reluctance has, until now, been supported by the conventional wisdom that subnational VATs are not technically feasible. Contrary to what has long been thought, however, such taxes are feasible. Subnational VATs may thus, over time, become the most important source of regional revenue in at least some of the larger federal countries in Latin America.

Third, admirable as the user charge and property tax package conventionally recommended for financing local governments is in many ways, experience has made it clear both that these proposals are very difficult to implement and that they are unlikely to provide an adequate fiscal base if local governments have major spending responsibilities in the social area. One result has been the proliferation of a variety of bad local (and regional)

³³In principle even subnational governments that depend heavily upon intergovernmental transfers may spend such funds efficiently and responsibly *provided* that all the conditions discussed in Section 2 are fully satisfied so that budget constraints are hard and local decisionmakers bear the full consequences of their decisions at the margin. Since few, if any, countries are likely to achieve such perfection, however, the better part of wisdom would seem to be to follow the advice in the text to the extent possible.

taxes on business. Partial relief from the resulting distortion may lie in the introduction of a “business value tax” (BVT) at a low and uniform rate. Even major reforms along these lines would not

solve all the problems of establishing sound and workable subnational tax regimes in all developing countries. Such reforms would, however, appear to move matters in the right direction.

Designing Intergovernmental Transfers

Regardless of the revenue sources made available to local governments, transfers from central to local governments will undoubtedly continue to constitute an important feature of the public finances of many countries. Transfers are needed if, for whatever reason, services must be provided by local governments that do not have the fiscal capacity to finance them at levels considered adequate, if there are externalities associated with the services in question, or if a country wishes to take inter-regional differences in needs into account. A well-designed system of intergovernmental transfers thus almost inevitably constitutes an essential component of any decentralization strategy.

Experience suggests that no simple, uniform pattern of transfers will be suitable for all circumstances. Since one size will not fit all, the first task of the fiscal tailor is to know as much as possible about the client. In particular, since transfers reflect closely the nature of a country's political system, their inherently political nature must be taken into account. One way to do so is simple, if somewhat artificial: focus on the *effects*, rather than on the *instruments* used to achieve them (Bird, 1993). Transfers as such are neither good nor bad. What matters are their effects on such policy outcomes as allocative efficiency, distributional equity, and macroeconomic stability. If, for example, the sole objective of fiscal decentralization is the efficient delivery of public services, as is assumed in the benefit model of Section 2, then all that matters is how transfers affect the effectiveness and efficiency of public sector operations.

The critical issue about intergovernmental transfers is, thus, not who gives them, or who gets them, or what the details of program design are, but solely their *effects* on policy objectives. As with user charges, the idea is to “get the prices right” in the public sector by designing transfers so that they do not weaken the hard budget constraint that exists if local citizens determine and

pay for the services that they receive (at least at the margin). Even if, as is likely to be the case in many countries, local democracy is not perfect or well-informed, it is still critical to ensure that transfers do not worsen outcomes by bailing out the incompetent and the irresponsible.

This focus on results suggests some desirable characteristics of transfer programs. Efficiency requires that those responsible for the provision of any service have a clear mandate, adequate resources (including, those they themselves raise), and sufficient flexibility to make decisions, while being held accountable for results. Transfers must be designed to satisfy these conditions. In addition to being “incentive-compatible” in this sense, simplicity, objectivity and transparency in transfer design are also important characteristics if the system is to work well.

Properly-designed transfers can achieve this goal even if they finance 90 percent of local expenditures. Poorly-designed transfers will not, even if they finance only 10 percent of expenditures. There is no analytical rationale for the argument sometimes made that accountability requires, say, half or more of local expenditures to be financed from local sources. However, it is obviously important to ensure that larger, more powerful jurisdictions—those where failure is likely to give rise to significant externalities and which are thus more likely to be bailed out—are self-financing to the greatest extent possible. Fortunately, these are also usually the places where potential local revenue bases are the greatest so that properly-designed transfer programs may have the greatest galvanizing effect on local revenue mobilization.

In reality, of course, many transfers are intended to achieve broader political goals, such as securing and maintaining stability either by rewarding friends or buying off enemies. Such objectives are by no means unimportant or irrelevant in many countries, as Breton (1996) has convincingly argued in general terms (and as Treisman

(1998) has recently demonstrated empirically for Russia). Nonetheless, from the economic perspective, it is important to minimize the collateral damage done to efficiency objectives by such “political” transfers.

Key Design Options

Three key factors in the design of intergovernmental fiscal transfers are the size of the “distributable pool,” the basis for distributing transfers, and conditionality.

Determining the Distributable Pool. Stability and flexibility are important characteristics of any good system of intergovernmental grants. How can these apparently contradictory characteristics be achieved simultaneously? Basically, there are only three ways to determine how much money is to be distributed through intergovernmental fiscal transfers: (1) as a fixed proportion of central government revenues or some other macro basis, for example, as a percentage of GDP; (2) on an ad hoc basis, that is, in the same way as any other budgetary expenditure; and (3) on a “formula-driven” basis—for instance, as a proportion of specific local expenditures or in relation to some general characteristics of the recipient jurisdictions. Variants of all these methods are found around the world.³⁴

If the central government wishes to maintain maximum political and budgetary control, the best system might seem to be one in which the total amount to be transferred (the “distributable pool” or the “primary distribution”) is determined annually in accordance with budgetary priorities. Under such a system, however, subnational governments will neither be able to budget properly nor will they face an appropriately hard budget constraint. Moreover, central governments that are susceptible to pressure from subnational governments may wish to insulate themselves to some extent by adopting a less discretionary system.

A better way to provide both some degree of stability to local governments and some degree of

³⁴See the country discussions in Shah (1994), Ahmad (1996), and Bird and Vaillancourt (1998).

flexibility to the central government is by establishing a fixed percentage of all central taxes (or current revenues) to be transferred. Sharing specific national taxes is less desirable than sharing all national taxes because over time central governments will understandably tend to increase more those taxes which they do not have to share. When national taxes are very sensitive to external shocks (such as a fall in export prices), however, this approach, although it may still be justified as sharing the pain between levels of government, may provide insufficient stability for the subnational finance of such vital human capital development services as education and health. For this reason, it may therefore be desirable either to base the total amount transferred on a more stable macro measure such as a moving average of GDP growth (as is done, for example, in Canada) or perhaps to finance such services separately through a system of capitation grants (Annex 6).

The Distributive Formula. A sound transfer system distributes funds among recipient jurisdictions on the basis of a formula. Discretionary or negotiated transfers are undesirable. Needs and capacity are the essential ingredients of most formulas for general transfer programs (as opposed to “matching grants” which are intended to finance narrowly-defined projects and activities—Annex 7). Needs may be roughly proxied by some combination of population and the type or category of municipality, or more refined measures might be used.³⁵ A more difficult, but conceptually critical, problem is to include a measure of the capacity of local governments to raise resources, given the revenue authority at their disposal. Even though, to remain transparent, formulas should not be too complex, any desired de-

³⁵An example of refined measurement is the Australian system sketched in Annex 5. Greater refinement need not mean better measures. Colombia, for example, includes a measure of “unsatisfied needs” in its transfer formula. The result, however, is that the formula assumes that it costs more to provide a given unit of service in one locality than in another, even though there is no reason to expect any particular association between needs and costs (World Bank, 1996a).

gree of interjurisdictional equalization can be build into such a formula (World Bank, 1996a).³⁶

A good transfer system, for example, might be structured to provide each local government with sufficient funds (own-source revenues plus transfers) to deliver a (centrally) predetermined level of services. Because capacity-based transfers are, in principle, based on measures of *potential* revenue-raising capacity (not on *actual* revenues), as developed below no disincentive to fiscal effort is created by this approach. Differentials in needs and in the cost of providing services (for example in rural or less densely-populated areas) may be taken into account although caution is necessary in introducing refinements since it is all too easy to turn a simple, transparent formula into an obscure and manipulable one. (Annex 5 outlines capacity-based transfer systems in Canada and Australia).

Conditionality. Once the total amount to be distributed has been decided and the basic distribution formula determined the question remaining is whether the transfer should be made conditional on the provision of certain services at specified levels. Two quite different approaches may be taken. On one hand, to the extent that the primary objective of the transfers is to ensure that all regions of the country have adequate resources to provide services at acceptable minimum standards, simple “lump-sum” transfers, with no conditionality other than the usual requirements for financial auditing, will suffice (Shah, 1994). This approach assumes that the fact that the funds flow to locally-responsible political bodies will ensure sufficient accountability and that it is neither necessary nor desirable for the central government to interfere with or influence local expenditure choices.

On the other hand, if the central government is, in effect, using local governments as agents in executing national policies (such as providing primary education at a specified level throughout the country), then it would make sense for the attain-

ment of some level of educational performance to be made a condition of the transfer (Bird, 1993). Since money is fungible, even transfers based solely on need do not ensure that the recipient governments will use the funds they receive as the central government might wish—unless receipt is conditioned on performance and compliance is monitored in some way. Although the extent to which fungibility is a problem depends upon whether the preferences of donors or recipients should rule, in the circumstances of most developing countries some conditionality often seems desirable, particularly when important national services such as education and health are provided by local governments (Annex 6).

Expenditure conditionality ensures that grant funds are spent on the specified service. It does not guarantee that grant funds do not simply displace “own” revenues that would otherwise have been spent upon the service in question. The extent of such substitution may vary greatly from service to service (Slack, 1980). *An alternative is performance conditionality*, which focuses on outputs rather than inputs (for example, the proportion of students achieving certain standards rather than the amount spent on education). This approach has considerable merit. For example, it focuses on what is presumably the real policy objective: education. It also, however, makes much greater demands on the ability of the central government to interpret the inevitably incomplete (and perhaps biased) information it receives. Experience suggests that, while it is useful to attempt to develop such indicators, it would be unrealistic to think that the many problems in doing so will soon be solved (Bird and Banta, 1999).

Fiscal Capacity and Fiscal Effort

Some measure of fiscal capacity is an essential element in the design of intergovernmental transfers to provide appropriate incentives both to local fiscal effort and to the appropriate use and management of transfer funds by recipient governments. In the absence of such a measure, a rough approximation may be attempted by using estimates of per capita regional GDP or some other macroeconomic index. Unfortunately, experience suggests that such indexes may not be

³⁶It is often argued that formulas should also include explicit measures of *effort*, but effort can be adequately taken into account in a capacity-based formula.

closely correlated with a reasonable tax capacity index (World Bank, 1996b). The “representative tax system” approach used in such countries as Canada and Australia (Annex 5) has not yet been tried in Latin America.

Although the evidence in support of the oft-asserted “fiscal laziness” of local governments and their failure to exert adequate effort to mobilize resources is far from overwhelming (Bird and Fiszbein, 1998), it is often argued that explicit measures of fiscal effort should be included in transfer formulas. There may indeed be reason for concern about the lack of will to utilize even the limited revenue mobilization discretion usually available to local governments. The solution, however, is unlikely to lie in playing with the terms of the distributive formula. Indeed, such a step is seldom advisable. Conceptually, fiscal effort may be most meaningfully understood as the ratio of actual taxes collected to “potential” taxes estimated on the basis of some standard measure of fiscal capacity and some standard tax rate (such as the national average). So defined, the absence of reliable empirical estimates of fiscal capacity renders the concept largely non-operational in most countries.

In any case, the measurement of fiscal effort is more complex than simply calculating such a ratio. If, for instance, tax bases are sensitive to tax rates, then this measure will overestimate capacity in areas with low tax rates (and hence underestimate the effort needed to increase tax rates) because the base will decline if the rate is increased. Moreover, given the limited flexibility of most subnational governments to alter revenues, it is far from clear the extent to which revenue behavior reflects “own effort”. Placing undue weight on fiscal effort in allocating grants may, for example, unduly penalize poorer areas, where, by definition, a given percentage increase in effort is more difficult to achieve. The problem giving rise to the need for equalization in the first place is that the fiscal capacity (tax base) of poor areas is too low, not that their tax rates are too low. Imposing an additional penalty on such localities is hard to justify.³⁷

³⁷On the other hand, experience in some countries sug-

In addition, including *actual* tax rates in any formula undesirably opens it to manipulation by recipient governments and hence distorts their allocative decisions. In fact, some such effects occur even in capacity-based grants when actual rates enter into determining a “standard” rate (Smart and Bird, 1997). Such problems will be even greater if the actual tax rate of any particular jurisdiction actually determines to some extent the size of the transfer it receives. Such arguments, combined with the fact that equalization transfers including a capacity factor embody a strong implicit incentive for transfer recipients to levy taxes at least at average levels, suggest that it is neither necessary nor desirable to include explicit effort factors in transfer formulas, even if such factors could be calculated in some reliable way.

Nonetheless, it *is* important to take fiscal effort into account in a more general sense in designing transfers (Wiesner, 1992). The reason is simply that local citizens should be required to pay in some meaningful sense for what they get so that those who make local expenditure decisions can be held accountable (through local political institutions) for their actions. So long as local governments are spending what they and their constituents view as “other people’s money,” they are unlikely to be under much local pressure to spend it efficiently. Experience shows that people are more careful in spending money they have to earn (taxes they have to pay themselves) because they are aware of the burden of taxation and because they will feel ownership of the activity as a result. Unless increased transfers are matched by a local

gests that introducing an effort correction into fiscal transfers may actually give *more* to poorer areas (that is, increase the redistributive effect of transfers) because poorer areas, once their much smaller fiscal capacity is taken into account, often levy relatively higher taxes than their richer neighbors. Again, it is by no means clear that this outcome is desirable. In practice, neither of these effects has been noticeable in most countries with “effort” measures in their formulas, either because the measure is seriously deficient (for example, the simple percentage increase in revenues) or because the weight attached to this factor is so small (World Bank, 1989, argues that both these problems existed with respect to the effort factor in the Colombian municipal grant).

contribution (however small that contribution may be in the poorest communities) the full efficiency benefits of decentralization are unlikely to be realized. People do not, it seems, take ownership of what is given to them in the same way as they do of things they have to pay for themselves, at least in part. And without local ownership, expenditure efficiency seems unlikely to be enhanced by decentralization.

Transfers and Poverty

Some analysts assess intergovernmental transfers in terms of how they affect the development of different regions. Others do so in terms of how successfully such transfers serve as proxies for direct interpersonal distributional measures. In Argentina, for example, although most transfers to provinces are unconditional, the relationship of transfers and poverty is broadly positive; poorer provinces receive more support from the central government, though it is not the very poorest that gain most. The regional redistributive process seldom runs smoothly, however. Porto and Sanguinetti (1993) show that per capita transfers to the poorest provinces (Chaco, Formosa, and Santiago del Estero), where 40 percent of the population lives in poverty, are only slightly higher than the average per capita transfer to all provinces. On the other hand, per capita transfers to some relatively wealthier provinces (such as Catamarca) were almost double the average.

Transfers, particularly those for investment purposes, are clearly influenced by political and territorial criteria. In Chile, for example, most regional investment is done by the central government. As in Argentina, the results favor southern frontier regions with low population density (Aisén and Magallanes). Nonetheless, per capita investment in Chile's poorest regions (Maule and Biobío) were close to the 1991 average (Espinoza and Marcel, 1993). Of course, many poor families live in rich regions. Over half of all Chileans with incomes below the poverty line live in the wealthiest regions of the country: the Metropolitan Region of Santiago, Valparaíso and Libertador (all in Central Chile). Territorially-based transfers are inherently inefficient in reaching these poor families.

The greater the weight the central government attaches to the equitable delivery of essential services to specific groups of poor citizens, the less the weight that can be given to local government autonomy. If a principal aim of central policy is to deliver "basic needs" to the poor, it should do it itself or develop an elaborate and detailed (and probably not very effective) monitoring system to ensure that local governments perform their role as agents of central policy efficiently and equitably. Decentralization, properly carried out, may have many virtues, but doing exactly what the central government wants in terms of income redistribution is not likely to be one of them.

Targeting the neediest is always difficult in developing countries. Nonetheless, Chilean municipalities appear to have been able to implement a fairly sophisticated method of targeting the poor. Municipalities maintain social welfare departments to gather, update, and verify data through household visits. Information from these questionnaires are fed into a national database, and national household surveys check municipal data periodically (World Bank, 1993). Chile's combination of better staff (see Section 7 below) and better data appears to allow it to come closer than most countries to achieving the level of sophistication needed at the local level to implement such specific poverty-alleviation schemes as direct delivery (or contracting) of preschool meals and nutrition supplements.³⁸

When local governments are expected to play a major role in delivering social services, they inevitably depend on central government transfers. Experience suggests that successful decentralization with a focus on poverty alleviation, must take care not to complicate intergovernmental finance unnecessarily (Bird and Rodríguez, 1999). Fiscal relations between governments are inevitably

³⁸A recent study in the much less developed country of Albania has similarly suggested that local targeting may sometimes be more effective than central efforts in relieving poverty (Alderman, 1998). Such outcomes assume either that local elites support these efforts or that the central government effectively monitors local efforts, or both.

complex. If the central government wants to accomplish a specific goal such as delivering specific services to specific (poor) households, it should do so, without further complicating inter-governmental finance. Many of the complications, and complaints, that characterize intergovernmental fiscal relations in most countries result from overloading the system with tasks for which it is ill-equipped, such as targeted poverty alleviation. Whenever feasible, direct transfers to the poor are better than indirect transfers to localities intended primarily to help poor households.³⁹

Similarly, if the provision of a nationwide basic uniform (or minimum) level of health or educational services is an important objective of national policy, the national government should either provide such services itself or directly transfer the needed resources to the target population, perhaps through some voucher system. If, for whatever reason, key poverty-related services *must* be provided through decentralized governments then (in addition to such issues as those discussed in Annex 6), very careful attention has to be paid to:

- getting the prices facing service providers right (for example, through a well-designed system of grants),
- setting up an information and inspection system able to ensure that the desired services are delivered to the target groups, and
- devising a system (such as a “fail-safe” clause for direct national provision) to deal with non-compliance without punishing the innocent.

These are not easy tasks. For example, experience suggests that it is not easy to develop an effectively decentralized health system that provides acceptable minimum levels of service in poor regions, let alone to poor people. In Chile, municipalities have for some years provided basic medical care, with most financing coming from earmarked subsidies from the central government. The subsidies were initially based on the number

and type of treatments provided by municipal health posts, and the level of reimbursement was intended to reflect the full cost of each specific type of treatment. The original open-ended reimbursement system, however, prompted a high level of demand and an unsustainable increase in total subsidies. As a result, in 1983 the central government imposed an arbitrary ceiling on these payments, but health expenditures continued to increase and had to be financed from local resources. From 1987 to 1990, municipal subsidies for primary health increased at an average annual rate of 61 percent, although they still amounted to only 4 percent of municipal expenditure in the latter year (World Bank, 1993). Similar problems have occurred in other countries with decentralized health delivery systems and centralized finance (for example, Canada).

Conclusion

Three distinct types of transfers play a role in most countries.⁴⁰ First, programs with strong national rationales and substantial interpersonal redistributive elements (such as education and health) might be separated out and funded through a national capitation system (Annex 6). Such programs are focused on redistribution among persons rather than regions. Central governments also have an interest in the provision of physical infrastructure related to these programs as well as other projects with significant spillovers. These investments might be financed on a matching grant (cofinancing) basis (Annex 7). However, these transfers will not work properly unless a final critical component of the system of intergovernmental transfers, which may be called a “general purpose” grant, is put into place. Such a grant has two important, and distinct, rationales. The first rationale, and the most important from a systemic perspective, is to provide the necessary underpinning for decentralization, and for the other transfers by equalizing (to some level) the *fiscal capacity* of departments and municipalities. The basic economic case for some form of general

³⁹ Indeed, such transfers may even have perverse effects. Localities that receive more money if they have more poor people have no incentive to reduce the number of poor.

⁴⁰ Although some or all of the features of these three types of transfers may be combined in a single instrument, it is nonetheless important to keep them analytically distinct.

transfer is to enable poorer local governments⁴¹ to respond adequately to central transfers intended to generate the appropriate level of public goods.⁴² A second rationale may be to provide sufficient

resources to enable all local governments, even the smallest and poorest, to provide a basic package of local services other than health, education, or infrastructure.⁴³

⁴¹Note that this means “poor” in terms of their capacity to raise resources out of local taxes imposed on local residents *not* necessarily in terms of how high the private incomes of those residents or the output of the locality may be. This distinction is important: local people can be rich, but the local government may nonetheless be poor because it has access only to a very limited range of taxes and charges.

⁴²The case for at least some unconditional general transfer support to most (perhaps all but the very richest) local governments thus has nothing to do with redistribution but is based on an efficiency argument. This need not be the whole story in efficiency terms. There may, for example, be growth-detering effects of such transfers on migration.

⁴³Although from a purely economic point of view, this objective may appear to make little sense, one rationale, similar to the familiar “basic needs” argument, might be that all citizens should be entitled to some basic level of local services, regardless of where they happen to live. Although the lack of local resources should not be confused with a lack of local capacity to make and implement suitable expenditure decisions (World Bank, 1995b), it is nonetheless evident in many countries that many poor local governments will always be dependent on transfers. A critical question in transfer design is therefore how to provide the basic resources such non-viable municipal governments need to survive while maintaining adequate incentives for them to do what they can in terms of raising their own revenues. Fortunately, it turns out that the same basic general support transfer design needed to make the capitation systems for education and health (Annex 6) and the matching grant system for infrastructure (Annex 7) can readily accommodate the politically necessary task of ensuring the survival of even the fiscally unfit.

Improving Local Capacity

One concern about the possible dangers of decentralization arises from the poor quality of local administration in most developing countries (Prud'homme, 1995). While there is often good reason for such concern, to considerable extent countries get the local governments they want and deserve. Subnational politicians and officials, like those at the central government level, respond to the incentives with which they are faced. If those incentives discourage initiative and reward inefficiency and even corruption, it should not be surprising to find corrupt and inefficient local governments. The appropriate response is to adjust the formal and latent incentive structures affecting local (and central) decisionmakers to make it possible and attractive for honest, well-trained people to embark in a career in local government. Given appropriate incentives (in terms of heightened expectations of improved services from their constituents and access to resources for which they are politically responsible) even very small local governments have demonstrated significant improvements in administrative capacity within a relatively short time (Fiszbein, 1997).

If expenditures, taxes, and transfers are set up properly along the lines sketched earlier, decentralized provision of services seems in principle at least, as likely as centralization to prove fiscally responsible. Moreover, it should be both more efficient and more accountable. Nonetheless, many are dubious of such assertions, not least because of experience in countries in which decentralization has not been well structured. Central politicians and officials are generally among the first to deplore the inevitable examples of corruption, malfeasance, and general incompetence that are likely to emerge in countries in which (often for the first time in their history) people at the local level are being asked, encouraged, and required to make decisions for themselves and to bear the consequences of them. Some of the people who feel this way may lose their jobs if they can be done more efficiently and in a more accountable fashion in a decentralized system.

More generally, like all major institutional changes, decentralization is a complex and lengthy process. Dillinger, Perry, and Webb (1999) recently noted that many of the alleged macroeconomic problems arising from decentralization are transitional. Much the same may be said of some of the implementation problems often experienced at the local level.

Those who emphasize the local capacity problem often suggest one of two solutions. If they support decentralization, they usually propose increased central support of local training and upgrading. If they are less supportive of decentralization, they often in effect take back the devolved power through means such as earmarking and mandating local expenditures (Annex 2). While there is little to be said for the latter approach, even the former may sometimes be slightly poisoned since there is no reason to believe that the best or only source of technical support and training for newly-decentralized activities should come from the central government.

During the last decade Colombia has made major strides toward decentralizing responsibilities and resources to its two levels of subnational government (*departamentos* and *municipios*). Considerable doubts have been expressed as to the effects of the newly decentralized system on the poor, particularly with respect to social spending. Both the willingness and the capacity of local governments to deliver effective assistance to alleviate poverty have been questioned (World Bank, 1994).

Although partial and preliminary, the evidence with respect to the development and use of local capacity in Colombia is surprisingly encouraging. A study of 16 municipalities across the country found numerous positive results in terms of enhancing local capacity (World Bank, 1995b).⁴⁴

⁴⁴As Maldonado (1999) emphasizes, however, this study sample was non-representative, since it was de-

Some municipalities, for example, improved the skills of local officials through competitive hiring, sharing the services of professionals among municipalities, training municipal employees, and rotating personnel across different departments in the same municipality. Capacity in terms of capital was also increased. One municipality privatized road maintenance; another put private developers in charge of the construction of urban roads. Computers have been introduced to monitor water and sanitation services in other localities. Some municipalities have started to share equipment. Some have also improved their technological capability (in terms of internal organization, planning, and monitoring) to ensure better management of municipal projects.

Underlying these improvements is a more basic change. In effect, some Colombian municipalities have moved to a “demand-driven” approach to public services as opposed to the previous “supply-driven” approach. Increasingly, reflecting both the liveliness of local politics and more extensive community participation, people are getting what they want, rather than what someone in the capital thinks they should want. Most municipalities have emphasized roads, education, and water works. These priorities may not always coincide with what the central officials previously in charge of these areas think is most important, but since these are the needs people perceive, these are the needs that the newly empowered and responsive local governments are attempting to satisfy. Opinion surveys suggest that the resulting sectoral allocation of resources is consistent with community preferences, with most respondents indicating that they trust the local government more than the national government to deliver goods and services (World Bank, 1995b).

As always, people are central to success, and the key to successful decentralization in any area is to provide the right incentives and institutional structure to induce and permit the right people to be in the right places. This is not always easy, and more attention should be paid to this critical “human” dimension of decentralization. Since newly empowered local governments are unlikely to de-

liberately chosen to reflect largely positive experiences.

liver services efficiently or to maintain (or meet) high quality standards, one way to deal with potential capacity problems has sometimes been to reassign central government workers to local governments. In the Philippines, for example, central government employees who were transferred to newly-decentralized local services in 1991 received on average around US\$142 a month (at the official exchange rate) compared to only US\$88 for local employees (Bird and Rodriguez, 1999). In these circumstances, improved quality and delivery of services may lead to an increase in local salary bills. Indeed, more generally as Campbell, Peterson, and Brakarz (1991) have argued with respect to Latin America, if service levels are to be maintained, decentralization may well increase rather than reduce government outlays, at least initially, through a “levelling-up” of salaries and service levels.⁴⁵ Over time, of course, cost-saving incentives may come into play and offset this initial effect.

A related problem that may arise concerns the location of central government officials who are reassigned to the local level. In principle, the most motivated public servants should be relocated to the areas of greater need. In most countries, however, most public employees work in the capital region, and there is generally no observable relation between regional “needs” and the ratio of public employees. A greater concentration of public employees in poor areas may not necessarily reflect a change for the better. Public bureaucracy is an essential part of the delivery of government services, but it may also reflect the creation of jobs for partisan supporters. Such a pattern seems clear in Argentina, for example, where the richest provinces (such as Buenos Aires and Cordoba) have less than 30 public workers per 1,000 residents while poorer provinces (such as Formosa, La Rioja, and Catamarca) have more than 80 per 1,000 (Porto and Sanguinetti, 1993).

⁴⁵Stein (1998) found that decentralization had indeed led to larger government outlays in Latin America, although the main reason appeared to be inappropriately designed intergovernmental fiscal relations which resulted in subnational governments not being subject to hard budget constraints.

On the other hand, both Colombia and Chile demonstrate the importance of qualified and involved local officials and the possibilities of improving local capacities in these respects. Municipal governments in Chile were once largely politically influenced employers of low-productivity labor. After more than a decade of reform, however, although the number of government workers per capita has dropped by 12 percent, the proportion of skilled professionals employed at the local level had increased (World Bank, 1993). An important starting point for this reform was a 1976 law establishing equivalence between municipal and central government employees in labor relations under the Public Sector Administrative Statute.⁴⁶ The Colombian study mentioned earlier found that some municipalities had made considerable progress in improving the level and profes-

sionalization of their staff (World Bank, 1995b). One municipality, for example, had channelled some of its unskilled laborers into a cooperative enterprise to maintain roads. Another had increased its level of "professionalization" through conscious personnel attrition policies. Evidence suggests that there are substantial economies of scale in this respect: a core minimum number of professionals is required for efficient local government, but the size of this core does not increase proportionally with population. Earlier Colombian experience had shown that there are a very wide variety of sources other than the central government from which localities can draw expertise. These include professional associations, universities, state enterprises, regional service agencies, and nonprofit organizations (World Bank, 1989).

⁴⁶Central regulation of local government employment may also exacerbate problems, however. In some countries, central control over employment (for example, in education) encompasses numbers, qualifications, and salaries with the result that virtually the entire local current expenditure budget (80 percent or more of which is wages as a rule) is predetermined from above (see Annex 2).

Macroeconomic Aspects of Decentralization

An interesting aspect of the recent move to decentralize public sector activities in many countries has been the revival of an old worry; namely, that subnational governments, left to their own devices, will act in a macroeconomically “perverse” fashion.⁴⁷ In Latin America, the fear is that increased national transfers will induce subnational governments to cut their own taxes while simultaneously expanding expenditures both through increased transfers and through “leveraging” increased borrowing on their new (transfer) revenue base. Subnational deficits, and hence total public sector spending and the overall public sector deficit, will thus expand. As a result, many countries developed new and more stringent controls over subnational borrowing (Ter-Minassian, 1996).

Transfers and Deficits

When a central government transfers revenues to subnational governments, as a rule it also transfers expenditures that are supposedly roughly equivalent. If so, subnational expenditures should, in the absence of behavioral reactions, increase by exactly the same amount as subnational revenues. Revenues available for expenditure at the central level decline and those available at the subnational level rise by the same amount, as do subnational expenditures. The macroeconomic impact of such a change depends not on the amount of the transfers per se but rather on what happens to central expenditures and what happens to non-transfer (“own”) subnational revenues.

Subnational revenues may decline as a result of increased transfers (“fiscal laziness,” decrease in fiscal effort, or “substitution effect” of transfers),

⁴⁷The literature on the alleged “fiscal perversity” of subnational governments from the point of view of macro stabilization dates back to the early Keynesian era, when it was argued that subnational governments would tend to accentuate recessions by cutting back on their expenditures exactly when the economy needed stimulus (Rafuse, 1965).

they may remain the same as in the absence of transfers (“flypaper effect”), or they may actually increase local own revenues (“stimulative effect” of grants). On the whole, the evidence, such as it is, suggests that at least up to now there is little or no evidence of a decrease in local fiscal effort in most countries (Burki et al., 1999). Macroeconomic problems arising as a result of increased transfers must then occur because the central government has not reduced its spending proportionately. Except when increased transfers are unaccompanied by increased expenditure responsibilities, the key to their impact on macroeconomic stability lies largely in the behavior of central (non-transfer) spending. Moreover, since local governments do not control monetary policy, even the most irresponsible local politicians can cause macroeconomic problems only when their actions are in some way validated by central authorities, either by the central bank (monetizing subnational deficits) or by the Ministry of Finance (bailing out subnational governments). Decentralization may complicate macroeconomic management, but it does not shift the final responsibility from the central government.

The common emphasis on the macroeconomic dangers of subnational borrowing, for example, is thus somewhat surprising. Since subnational governments cannot print money, a systemic problem can arise only if one or more of three conditions holds:

- the central government validates subnational deficits by expanding the national deficit to finance them;
- the central government explicitly or implicitly subsidizes subnational borrowing (for example, by guaranteeing it); or
- subnational governments borrow abroad and spend the money in ways that do not add to the country's capacity to service foreign debt.⁴⁸

⁴⁸As familiar from the classic discussion of the “transfer problem” by Keynes and others, this is essentially

The first two of these conditions are clearly within the control of the central government. If it wants to avoid macro problems arising from subnational debt it can do so by not subsidizing such borrowing and by letting subnational governments that borrow too much go bankrupt—although the mechanics of how this might be done are by no means simple (World Bank, 1996b). If intergovernmental fiscal relations are properly structured, many problems can be avoided.

Subnational Borrowing

Nonetheless, there have recently been some well-known examples of subnational deficits and debt in practice, particularly in Argentina and Brazil (Dillinger and Webb, 1999). Similar concerns have arisen in other countries such as Mexico and Colombia. Given the prevalence of bailouts and discretionary transfers in such countries, it is clear that there is reason to be concerned about this problem. In effect, central governments in most countries do seem to be providing implicit (if not explicit) guarantees to subnational borrowing. In these circumstances, they should indeed keep the door to the treasury locked to obviate the possibility of local profligacy and hence demands on central funds. Such practices are common even in developed countries. In Canada, for example, although there are no restrictions on *provincial* borrowing, the provinces themselves severely restrict *local* government borrowing in a number of ways: the amount of debt, the type of debt instrument, the length of term, the rate of interest, and the use of debt funds, are all, as a rule, strictly controlled.⁴⁹

what happened in Germany in the 1920s, as municipal governments borrowed abroad in order to finance such things as swimming pools.

⁴⁹Some provinces require provincial government approval before debt is issued; others require the specific approval of local electors. Sometimes the restrictions are different for different categories of municipalities or for short-term as opposed to long-term debt. But in no case are local governments allowed to borrow as they wish. For a detailed discussion of subnational borrowing at both the provincial and local levels in Canada, see Bird and Tassonyi (1999).

It has been proposed that (Artana et al., 1995) to deal with this problem, a strict limit on the borrowing ability of subnational governments should be put in place. The rationales behind such recommendations appear to be twofold. The first is simply that unchecked subnational governments, particularly those that are highly dependent on national transfers, are inherently likely to behave irresponsibly by increasing current expenditures well above their capacity to finance them out of current revenues and then closing the gap through borrowing (especially if, as in Argentina, they control provincial banks), often pledging future transfer revenues to service the loans.

To counter this, it might be argued that the subnational deficits observed in some countries may, to a considerable extent, have been a transitory rather than systemic phenomenon, reflecting the inevitably lagged adjustment of fiscal institutions after the irresponsible budgeting practices that developed during the hyperinflation years (Burki et al., 1999). In Argentina, for example, the rapid and unexpected rise in provincial revenues in the early 1990s during the first phase of fiscal adjustment led to some unwise expenditure increases that could not be sustained. Such asymmetric behavior is common. The solution is simple in principle, although seldom put into practice: mute the effects of the boom, and subsequent bust, by smoothing out fiscal effects (for example, through a stabilization fund). In addition, as always, try to improve public understanding of fiscal realities by such measures as increasing the transparency of financial reporting.

Principled solutions like these are seldom found in practice, however. Their absence may be symptomatic of a more basic underlying problem; namely, the general inability of central governments to credibly commit to a “no bailout” policy. In the absence of hard budget constraints local governments will have some incentive to misbehave. But it may not be possible in the circumstances of many countries to institute constraints hard enough to prevent this outcome. Some subnational governments may be too economically significant or too politically important to let them fail (Wildasin, 1997). In other instances, the effects of failure on local residents may be so unde-

sirable that the central government has no choice but to interfere. Indeed, no matter how well inter-governmental fiscal relations are designed there is likely to remain a serious short-term problem with respect to those subnational governments, which fall into the deficit trap. Emergency support from the central government may sometimes be needed to resolve such problems.⁵⁰ If so, however, assistance should carry with it the obligation to introduce effective reforms under the supervision of a review board. More general restrictions on subnational borrowing may also be advisable.

The second rationale for imposing special controls on subnational borrowing is that the central government must have full control over all the policy instruments needed to ensure macroeconomic stabilization. In particular, subnational governments should be restrained from accentuating cyclical pressures by going into debt. In particular, subnational governments should be discouraged from borrowing overseas and adding to the foreign debt. This argument suggests that transfers should be fixed amounts in any particular fiscal year (Artana et al., 1995). In addition, in countries such as Argentina where provincial control of financial institutions facilitated excessive borrowing and its concealment, privatizing provincial banks and making provincial fiscal behavior more transparent through improved accounting and reporting procedures is also desirable.

Subnational governments may borrow for many reasons, some more defensible than others. In some cases, for example, local officials have reaped the political benefits of expenditure financed by borrowing, while leaving the political pain of debt service to their successors. Such

⁵⁰Note that this does not necessarily imply subsidization. In Canada, for example, the full cost of any needed financial restructuring is supposed to be borne locally in the form of reduced services for a period of years until the debt is cleared. The rule is clear: “your mistake, you pay.” The central (provincial) government supervises the review and adjustment process, but it does not subsidize it except very occasionally when the default was clearly beyond the control of the locality (for example, a forest fire that destroys the resource base of a rural community) so that the problem is one of solvency not liquidity.

practices raise a major issue of accountability. A policy that might be considered to stem this behavior is to prohibit government officials from raising their own salaries. Legislatures could vote salary increases for themselves and for local officials, but no increases could take effect until after the next election.

In other instances, subnational governments have been forced to borrow (even illegally) because they are so constrained by central rules that the only resources that they can freely allocate at the margin in response to constituent demand are borrowed funds. Is it not surprising that some governments choose to increase borrowing when after years of rapid increases, transfers level off. From the point of view of strict fiscal virtue they should have reduced expenditures at once. But even local politicians are human and tend to take the easy way out, if they can find one.

Finally, borrowing may be the economically appropriate way to finance capital outlays. As every public finance textbook notes, in terms of both allocative efficiency and intergenerational equity it often makes sense to finance long-lived investment projects by borrowing rather than relying upon current public savings or transfers.

Despite the potential for subnational fiscal indiscipline, there are some instances when borrowing is good. In principle, it is as important to facilitate “good borrowing” as to block “bad borrowing.” Indeed, it may even be more important since, in the right conditions, the market may impose the needed discipline. From this perspective, imposing too strict and arbitrary central limits to subnational borrowing may have a perverse result. Like deposit insurance, debt limits and similar controls raise moral hazard problems precisely because they prevent market discipline from being applied. Potential lenders to subnational governments should reasonably be expected to be sufficiently capable and motivated to find out what risks they face. If such loans enjoy the approval of the central government, they are relieved of this task and also provided with a stronger implicit guarantee that the central government will back it up. The result may be more risky lending.

For a hard budget constraint to be effectively imposed by capital markets, there must not only be a credible no bailout rule, but in addition there must also be transparency so that lenders have full information on borrowers and local residents have full information on the consequences for them of local borrowing and the ability to influence local decisionmakers. To the extent that democracy and markets work together in bringing about responsible fiscal behavior, the process is likely to take time. Perhaps, as has been argued with respect to some developed countries, the “hardest” budget constraints are those forged in the fires of experience (Bird and Tassonyi, 1999). If countries have learned through experience that careful attention needs to be paid to subnational borrowing to avoid serious problems, then such problems are more likely to be avoided — whether by re-designing intergovernmental fiscal relations to reduce the temptation to borrow irresponsibly or by establishing a set of rules or signals constraining such borrowing. In the end, what may matter more than the specific solution adopted is the clear recognition of the problem and of the consequences if it is not dealt with properly. Most mature democracies went through decades, even centuries, of difficulty before working out the ways in which they presently cope with the potential problems arising from relatively autonomous subnational governments.⁵¹ Perhaps Latin American countries can shorten the process by learning from the experience of others, but in the end each country is likely to follow its own path with some having an easier time of it than others.⁵²

To some extent, the oft-expressed concern about irresponsible subnational governments getting themselves into trouble looks like yet another in-

⁵¹See, for example, the discussions of various developed countries in Ter-Minassian (1997); see also Rodden (1999), Inman (1999), and Bird and Tassonyi (1999).

⁵²For example, as Dillinger and Webb (1999) note, Argentina should be able to deal with this issue more easily than Brazil and, as the same authors note in another study, Colombia's conservative fiscal traditions may help it to cope with such difficulties in its own way (Dillinger and Webb, 1998).

stance of inappropriate and misconceived paternalism, a common attitude of central governments faced with the uncomfortable prospect of losing control as a result of decentralization. Subnational governments seem unlikely to be well managed if those in charge know that central government bailouts will save them from the consequences of their actions. Clear rules and limits on subnational borrowing will also save the central government from the consequences of its own inability to refrain from bailing out subnational governments. In other words, one reason for imposing restraints on subnational borrowing may be to restrain the central government from its own weakness (Elster, 1984). If central governments are all too likely to bail out defaulters, they have good reason to try to limit the extent to which this weakness can be exploited by limiting the possibility of default.

Although allowing subnational access to capital markets should, in the long run, strengthen rather than weaken fiscal discipline, the “long run” may be quite long. In the interim, it may be sensible to limit such access to some extent, for example, by introducing something like the Colombian system of “traffic lights,” under which subnational governments can borrow freely (green light) when debt is below a specified threshold, but require central approval (yellow light) when debt levels are higher (Dillinger and Webb, 1998). Given that formal bankruptcy proceedings in the public sector are impossible, reducing the moral hazard requires the institution of a credible review and control systems for working out debt, perhaps along the lines suggested recently for Argentina (World Bank, 1996b).

Subnational government borrowing in foreign markets may give rise to special problems. A sensible precaution might, therefore, be to require explicit prior approval from the central government before any such loans are entered into. More generally, to ensure accountability, all subnational borrowing should be reported immediately and in a transparent fashion so that no government can shift hidden debts onto the next administration and local voters and the central government can understand more clearly what is going on. The problems in Argentina, for example,

arose under a clearly inadequate set of rules with respect to accounting, reporting, and transparency. Such problems can, and should, be corrected. Moreover, since the only good case for local borrowing is to finance capital investment, no borrowing should be permitted for other purposes, no matter how worthy.⁵³

Finally, although some have expressed concern about the ability of local governments to borrow on the basis of the increased cash flow as a result of transfers (or, for that matter, royalties), it is not clear why this should be a problem if such borrowing is not subsidized. To the extent that transfers constitute revenues that local governments can spend as they wish, if private agents are willing to lend money based on this security they should be free to do so, provided that they also bear the consequences if the loan goes bad.

To sum up: as a rule, inappropriate borrowing by subnational governments should be viewed not as a problem of decentralization but as a symptom of more basic underlying inadequacies with the intergovernmental fiscal system in general. Once this is resolved (through measures such as the reassignment of revenues (and perhaps expenditures); a revised transfer system; the introduction of transparent, timely, and reliable reporting systems; and the establishment of a stable, accepted periodic review process,) the institutional problems that may give rise to unsustainable subnational borrowing should largely be solved. However, it may take considerable period of learning before practice reaches the same level as theory and, in the interim, certain specific rules and limits may need to be put into place to reduce the likelihood of undesirable outcomes.

Two distinct problems now face some countries (Argentina, Brazil) and may soon face others (Mexico, Colombia). The first is to restructure the finances of a few subnational governments that face difficult debt situations. The second is to restructure the system of intergovernmental finance in such a way as to ensure that the delivery

of services provided at the subnational level should be as equitable, efficient, effective, and stable as possible. These are essentially unrelated problems, and need to be dealt with separately. If the long-term problem is properly solved, however, the short-term problem should not arise again.

In the end, only two basic *ex ante* limits on subnational borrowing seem strictly desirable in most countries: First, borrowing should be permitted only for investment purposes (a restriction that may not always be easy to enforce in the absence of strictly segregated capital budgets). Second, explicit national approval should be required for borrowing abroad. At a deeper level, however, given the extreme dependence of subnational governments on central transfers (as well as political reality in many countries), somewhat stronger borrowing restrictions than these may be warranted.

Unless subnational governments are able to save themselves from fiscal crises by drawing on their taxing powers, their only options are bankruptcy or bailouts. The *fiscal* root of this problem is the limited taxing powers available to subnational governments that are expected and required to carry out a much wider range of functions than they can finance on their own without extensive reliance on central support—either directly through transfers or, less desirably, indirectly through bailouts (Eichengreen and von Hagen, 1995). The *political* root of the problem, however, lies in the continuing expectation by all players (citizens, subnational and national politicians, and lenders) that, in the end, the central government will come to the rescue. As long as central actions, *ex post*, reinforce this expectation, such common *ex ante* administrative controls on borrowing as requiring prior central approval or limiting debt service to a certain proportion of current revenues may have to remain in place.⁵⁴

⁵³Some arrangements may have to be made to permit “in-period” borrowing to smooth out cash flows over the budgetary year.

⁵⁴For an example of a phased (and assymmetric) approach to liberalizing subnational access to capital markets, see Ahmad (1999).

Information and Accountability

As has been emphasized throughout this paper, if decentralization is to work properly, providers of local infrastructure and services must be accountable both to those who pay for the services and to those who benefit from them —two groups that are seldom identical. Enforcing accountability at the local level is not easy. It requires the establishment of an intergovernmental fiscal system that provides clear and correct incentives to all relevant decisionmakers. It also requires the provision of adequate information to local constituents, as well as the opportunity for them to exercise some real influence or control over the service delivery system. “Informal” organizations almost by definition must be structured like this or they are seldom sustainable (de Soto, 1989). In the political and social circumstances of many developing countries, however, it can be a considerable challenge to introduce a similar degree of responsiveness into formal governmental organizations.

In any case, improved accountability is the key to improved public sector performance and improved information is often the key to accountability. The systematic collection, analysis, and reporting of information that can be used to verify compliance with goals and to assist in making future decisions is critical to successful decentralization. Such information is essential both to informed public participation through the political process and for monitoring local activity by central agencies responsible for supervising and (usually) financing such activity. Unless local “publics” are aware of what is done, how well it is done, how much it cost, and who paid for it, no local constituency for effective government can be created. Unless central agencies monitor and evaluate local performance, there can be no assurance that functions of national importance will be adequately performed once they have been decentralized, although care must be taken to avoid the use of such information for inappropriate control purposes.

An important accompaniment of any successful decentralization program is thus an improvement in national evaluation capacity. Decentralization and improved central evaluation and assessment of local activities are not substitutes; they are complements. An essential element of the hard budget constraint needed to induce efficient local decisions is adequate central enforcement capacity in the shape of credible information gathering and evaluation. The “carrot” of central financial support of local efforts must in the end be accompanied by the “stick” of some degree of withdrawn support if performance is inadequate. A credible stick requires both some standard of adequacy and some way of measuring performance.

Unfortunately, a common problem is the lack of an appropriate structure to monitor and support local governments. Central governments should, for example, be responsible for monitoring and assessing subnational governments finances, both in total and individually. Yet, most central authorities do not have a very good understanding of the situation faced by local governments or of the likely effects of any proposed changes in local finance. Most need much more regular and detailed financial data on local governments than they usually possess.

Ideally, uniform financial reporting and budgeting systems should be established, perhaps with different degrees of complexity for different categories of local governments. As discussed below, an appropriate agency (preferably one with a degree of political separation from the central government) should be made responsible for collecting and processing data in a timely fashion.

Central government support may be needed not only to develop and maintain an appropriate reporting system, but also to train and support local government officials. As noted in Section 7, if decentralization is to be effective, support is often needed to upgrade the technical capacity of local governments to carry out efficiently and effec-

tively the expenditure functions for which they are responsible. In some instances the necessary support might come from the relevant central ministry, while in other instances it may come from the ministry, regional agencies, universities, private sources, or some mixture of these bodies. Each functional area has different requirements, problems, and possibilities and needs be treated differently.

Different countries may have different cultural understandings and political conditions with respect to the desirable level and nature of accountability. Whatever standards and practices of accountability may be considered desirable, however, formal reporting and evaluation systems inevitably constitute essential components of any workable accountability system (whether to users, to local taxpayers, or to the central government, depending upon the source of financing). In Wisner's words (1993) "the nature and source of financing is the crucial determinant of efficiency." In all cases, an adequate system of collecting and assessing information is required not just for accountability but also, more importantly, to help establish a "public" to whom to be accountable.

The role of the central government in ensuring and monitoring effective decentralization is especially critical when, as is often the case in practice, a major concern is to enhance service delivery, particularly with respect to services such as health and education that are important determinants of national development, poverty and welfare. An unfortunate side effect of decentralization in some countries has been the virtual disappearance from the central government's cognitive horizon of reliable information on the provision of such services. Decentralization of a function does not mean that the central government no longer has any responsibility in the area. What it means is simply that the nature of central responsibility has changed from direct service delivery to regulating and monitoring the efficiency and equity of services delivered by local governments. The essential tool needed for this task is an adequate and up-to-date information base (generated, for example, by requiring local governments to file uniform and informative budgets and financial and other reports). Unfortunately, this need for ex-

tended and reliable information has been sadly neglected in most decentralizing countries, in which those making the decisions often understandably wish to hide the real reasons for (and the results of) their actions and to live the quiet life of the unaccountable.

An incentive to comply with demands for such information may be created by making timely submission of such reports a condition of receiving fiscal transfers. Implementing this requirement may in turn imply significant initial set-up costs in the form of system design, training, and implementation. To spend money wisely in a decentralized system may sometimes require some initial investment. While such concerns are less important when full responsibility for the efficiency and equity of services is devolved (as in a formal federal structure) rather than simply delegated to subnational governments (see Section 2 above), the need for an adequate information system remains critical to effective macroeconomic management (see Section 8), given the importance of such services as education and health in total public expenditure. Regardless of the form of decentralization, an important institutional problem is thus how to ensure *both* that the relevant central government agencies have adequate incentives to monitor subnational activity *and* that subnational governments have sufficient incentives to provide the necessary information.

Institutionalizing Decentralization

Decentralization is a dynamic process. No country ever gets it right on its first try. Circumstances change, and the nature and design of intergovernmental fiscal relations should change also. An important aspect of establishing an adequate institutional framework for decentralization is thus to build in some "error-correction" mechanism, that is, to permit and encourage the adaptive development and evolution of the system in response to changes in needs and capacities.

Several possible mechanisms along these lines are suggested by experience around the world. One approach, is to build in "sunset" provisions into transfer programs, that is, transfers to local institutions are subject to renewal in a specified num-

ber of years, provided they pass an independent performance evaluation. Another approach is to use the likely need for some centrally-supported access to capital markets for infrastructure finance not only as a screening device to reject obviously flawed projects (Annex 7), but also as an evaluation system to build up “ratings” of local capacity and effort. Yet another approach, emphasized in this section, is to assemble and publicize reliable comparative information on local government performance. Finally, as discussed further below, most democratic federal countries have developed specialized institutions that serve to integrate the fiscal government decisions at different levels and to provide the public with some useful and trustworthy (nonpartisan) information on what both levels of government are doing, separately and together.

Developing such institutional “infrastructure” is neither a quick nor a cheap undertaking, but it is essential if countries wish to decentralize public sector activities without losing touch with what is going on in these important parts of the public sector.

Much more informed and open discussion of these matters than now prevails in most countries is needed. Regular publication of relevant data would help, but one cannot rely solely on an interested party (the central government) to carry out, let alone publish, all the information needed for public discussion. In countries in which intergovernmental fiscal issues are important, consideration might therefore be given to establishing a small nongovernmental research institute focusing on local government problems.

Some central governments are learning through experience to carry out this essential information and monitoring activity more adequately in their own interests, for example, to avoid excessive subnational borrowing and hence potentially undesirable macroeconomic pressure. Nowhere, however, has the central government yet admitted that full disclosure (transparent accountability) should apply also to its own actions as they affect intergovernmental fiscal relations. Decentralization is a two-sided street, and the role played by the central governments in determining the out-

come of this complex process needs more attention than it seems to have received. As Breton (1996) has demonstrated, for the potential benefits of increased governmental competition to be realized, governments at *all* levels must become more transparently comparable and hence accountable for their actions.

Without a public better informed of the real alternatives and choices facing them in their collective capacity as members of society, the long-run sustainability and outcome of decentralization is questionable in many countries. The decentralized governments that now exist in most countries seem unlikely to be particularly efficient. But they can, and may, become more so over time if sufficient attention is paid to developing the institutional infrastructure needed to produce such a favorable outcome.

Achieving Fiscal Transparency

Fiscal transparency is fundamental to sound public policy. It is needed not only to improve the working of the executive and legislative branches of government but also to improve the level of public discussion and understanding of policy issues. The capacity to accept and absorb policy change in the public at large needs to be strengthened in countries in which, all too often, people have already suffered much from change while they have not yet absorbed such basic economics lessons as that one cannot get something for nothing and that life is not a zero-sum game. In general, the more open and transparent the public policy process is, the more likely are policy decisions to be grounded in fact rather than fantasy, and the more policy outcomes should coincide with stated policy intentions. Such arguments are now well accepted in theory, and to some extent supported by empirical studies and experience in a variety of countries (Kopits and Craig, 1998).

Nonetheless, fiscal transparency is inherently a difficult concept for many politicians and officials to accept. For them, there are often overwhelming incentives to hide taxes and liabilities, to exaggerate benefits, and generally to present fiscal realities in as complex and confusing a way as possible in order to present a more acceptable public

face. In this way, they can then live the quiet life available to those who exert monopoly control over information that might be used to make them accountable for their actions (or lack thereof).

The need for transparency and more informed public input, is nowhere greater than when the responsibility for governance is divided between two (or more) tiers or levels of government. Not only do citizens in such countries need to know exactly who is responsible for precisely what if they are to hold governments accountable through the electoral process (or in other ways), but it is also critically important that each level of government be well informed of the intentions, actions, and outcomes of the other levels. Although most apparent in formal federations, similar pressures, and needs, arise in all countries with some degree of fiscal decentralization.

As stressed earlier, the key to effective accountability, and hence to good governance, is to make relevant information publicly available in a comprehensive and comparable form. Uncomfortable though the lives of politicians and officials may be when their actions are open to public scrutiny, the more that is known, and the more publicly it is known, the better policy outcomes are likely to be over the long run. To achieve efficiency in a decentralized setting, the rules of the game must be clear and all players must have access to broadly the same information base.

Perhaps the main conclusion one can draw from international experience with specialized “decentralization institutions” intended to improve the nature and understanding of the process is that each country deals with these matters in accordance with its own genius (or stupidity). Nonetheless, experience around the world (Annex 8) suggests that several different roles might be envisaged for a body concerned with intergovernmental finance issues. Depending upon what is politically desired and feasible, some of these roles may, or may not, be incompatible and may or may not be combined in the same institution.

One of the most important institutional requirements for effective decentralization is a cadre of fiscal and policy analysts at all levels of govern-

ment. These persons need to be trained not only in the mechanics of analysis but in working together, from their different perspectives, the common goal of making the complex political and administrative system work. A possible role for a special intergovernmental commission or agency, for example, might be to serve as a link between the political and executive branches in the center and the regional (and perhaps local) governments. A special institution might achieve this aim by providing a nonpartisan forum within which various actors may get to know one another in a non-confrontational setting; by providing a common informational basis to all parties that is trusted by all parties; or by training both formally and informally a cadre of experts in fiscal and financial areas.

Whether the “forum” role (which would appear to require that members be selected largely on political criteria) is really compatible with the other two attributes mentioned is by no means clear, however.

An intergovernmental fiscal agency might support and develop subnational administrative capabilities in financial and fiscal areas, working with local institutions to develop and offer appropriate training and even helping to implement uniform statistical information systems. Such an agency might help develop the needed pool of professional financial analysts committed to public sector work in a variety of ways, through, for example, offering short training programs, affording promising young graduates internship positions and so on. Experience with such training programs in a variety of circumstances suggests that one of their major intangible benefits is the facilitation of frequent (and productive) interchanges not just between federal and regional actors but also between regional actors themselves.

Other ways in which an intergovernmental agency might, if desired, serve a formal political role might be by making annual or periodic reports on “the state of the federation”; by determining, or reporting on, the appropriate basis for grants; by providing “federalism impact statements” indicating clearly the impact on subnational governments of central actions; by working with central

banks and departments of finance in developing comparable public finance data (for example, on borrowing); or by going further and, like the Australian Loan Council (see Annex 8), serving as a formal coordinating mechanism for public sector borrowing.

The substantive content of such tasks might best be achieved by a body whose main formal role was educational rather than political. For example, a common accounting framework could be developed and utilized for monitoring the fiscal performance of all levels of subnational government. Such reports might be more credible if the agency did not have a formal “reporting” role to the central government, and its output was accepted as the work of competent and politically neutral analysts.⁵⁵

An intergovernmental fiscal body might also play a useful role from the point of view of the subnational governments by monitoring and reporting on the possibly intrusive impact of central government policies on subnational governments. Such work might be more credible if it were not seen as that of one of the players in the political game but rather as an outside, but well-informed and trustworthy, referee.

In the hurly-burly of emerging democratic politics in many Latin American countries, such suggestions may seem either naive or hopelessly idealistic.⁵⁶ Nonetheless, for a decentralized system to work well in a democratic setting, people need to

understand not only what is going on but also the real possibilities and constraints facing governments at all levels. If an appropriate agency can be established, it might serve as a central analytical focus, supplying trustworthy information to all actors and also helping to create the informed public that is needed for the political system to cope with the great strains and challenges facing it. How “political” such a body should be (like the details of its appropriate structure and functions) can only be determined in the context of each country.

Experience in developed federal countries suggests that transparency is needed for good fiscal management, that good fiscal management is needed for a successful decentralized structure, and that an agency performing at least some of the functions sketched above may play a vital role in this process. Good, relevant, timely analysis of intergovernmental fiscal relations, training good analysts, facilitating and encouraging productive technical exchange between and within governments, and providing neutral, competent input to the public discussion of intergovernmental fiscal and financial policy (such intangible but critical institutional factors) play a vitally important role in making decentralization work. If the current wave of decentralization in Latin America is to prove sustainable and beneficial in the long run, steps could usefully be taken to foster the development of such an institutional framework, certainly in the larger countries of the region.

⁵⁵On the other hand, such an agency might have more public impact if legislative bodies were legally required to receive and discuss its reports.

⁵⁶For an accurate, and pessimistic, assessment of Latin America's record to date along these lines, see Rojas (1999).

Conclusion

This paper has covered a very wide range of topics. One of its principal arguments is that many aspects of intergovernmental fiscal relations can be worked out specifically only in the context of each individual country. The extent to which any “ideal” structure of intergovernmental fiscal relations may be approximated in the case of any particular country depends upon constitutional restrictions, political realities and an array of specific contextual factors. Allowing for differing degrees of democracy or differing levels of competence would introduce more variations, as would allowing for different sizes and capabilities of subnational governments. Although such complexities cannot be encompassed in a simple tabular presentation, it may nonetheless be useful to summarize in box 1 the principal recommen-

dations with respect to intergovernmental fiscal relations that emerge from the extended discussion in this paper.

While most of the items listed in this table relate to such universal principles as establishing a transparent and accountable system (and hence apply in principle to all countries), several may be differentiated for types of countries (for example, (1) a large federal country in which important social services are provided by subnational governments, (2) a decentralized unitary country in which social services are provided at the subnational level but where the central government has more control over subnational spending, (3) a small unitary country in which local governments provide only “local” services).

BOX 1

Some Guidelines on Intergovernmental Fiscal Relations

Expenditures

- Clarity of assignment
- No unfunded mandates
- Only rational earmarking
- Different fiscal years
- Uniform framework laws (possibly differentiated by size and level)
- No prior budget approval
- External audit
- Public reporting
- Medium-term expenditure frameworks at all levels
- Subnational must have ability to cut expenditures
- Subnational must have ability to raise expenditures (financed by self)

Revenues

- User charges whenever feasible
- Residential property taxes (at local level)
- Ability to set tax rates
- Surcharge on VAT (for regional governments)*#
- Surcharge on PIT*
- Floor and ceiling on any business taxes (including nonresidential property tax)
- Selected excise taxes (vehicles, etc.)*

Transfers

- Total determined in stable/flexible manner (% GDP or revenues)
- Distributed by simple formula, including revenue “capacity” element
- No discretionary transfers
- Basic transfers can be spent as recipient wishes
- Capitation transfers for key social services**
- Matching transfers for infrastructure
- Periodic review of system

Borrowing

- No bailouts
- Specific “work-out” system to deal with liquidity crises
- Only permitted for investment projects
- Interim “traffic light” system of debt limits requiring approval
- Coordination of foreign borrowing

Institutions

- Framework laws (possibly differentiated by size and type)
- Support for training and capacity-building
- Support for nongovernmental analysis
- Intergovernmental agency for periodic review of system
- Central government “decentralization” agency##

Notes: *Regional governments only.

#In “federal” state, regional governments may levy variable surcharges; in “decentralized unitary” state, or where regional governments are small, it might be better to impose a uniform surcharge, with proceeds distributed by estimated consumption.

**This approach may be used in both decentralized unitary countries and in those federal countries in which it is constitutionally permissible.

##In a federal country, this agency would have primarily an informational and policy role; in other countries, it may play a monitoring and to some extent control role.

Annex 1

Expenditure Assignment in Chile

Chile, although a highly centralized country, provides a good illustration of both the use of the efficiency rule for expenditure assignment and its limitations from an accountability perspective. Chile has for some decades systematically undertaken the strengthening of local government institutions and resources. Care has been taken to define clearly the responsibilities of different levels of government. Municipalities are in charge of paving urban streets and rural feeder roads, drainage, solid waste collection and disposal, street lighting, open spaces, recreation, and cemeteries. Primary and secondary education and basic health care services are also delivered by municipal governments, but the central government is largely responsible for financing these services and ensuring that minimum standards are provided regardless of the poverty of local governments. Higher education and higher-level health care is the full responsibility of the central government, and, in contrast to many countries, the central government also provides and finances public safety.

On the whole, the level of institutional capacity at the local level in Chile appears to have been sufficiently high and responsive to local needs and demands to carry out the assigned functions adequately. Nonetheless, Chile still faces significant challenges in terms of improving local (and regional) revenue capacity and accountability. For example, the responsibility imposed on local governments in terms of expenditure assignment has not been matched by devolving sufficient

authority to them to do the job correctly. Although following the constitutional reform of 1991 local political democracy was restored with the elections of mayors and councilors, little has been done to give local governments greater discretion in levying and collecting taxes and charges. Local governments have only a very limited ability to raise taxes to meet residents' demands for additional services. Local governments with the capacity to do so are, to some extent, discouraged because under the present transfer system a share of any additional tax revenues has to be shared with other municipalities.⁵⁷

Nonetheless, on the whole, Chile's experience with decentralization appears to have been fairly positive. Considerable use has, for example, been made of such promising institutional innovations as contracting services. Over 70% of the municipalities contract out solid waste collection and disposal as well as maintenance of street lights (World Bank, 1993). As noted elsewhere, what matters with respect to local public services is *what* is delivered, not *who* delivers it. Experience in Chile, as in other countries, demonstrates that much of what is commonly thought of as traditional local government activities can, in fact, be carried out more efficiently in nontraditional ways (UNCHS, 1996).

⁵⁷Uniquely in Latin America, Chile has a well-developed "horizontal" revenue-sharing system (not unlike that in Denmark). While this approach has considerable merits in theory (for example, the "externality" argument for intergovernmental transfers suggest that they should be between governments operating at the same jurisdictional level) in practice it is difficult to relate the actual design of the *Fondo Común* to such considerations, and the disincentive effects of "taxing" those who raise more revenues suggest that this system should perhaps be reconsidered.

Annex 2

Earmarking and Expenditure Mandates in Colombia

Colombian public finances have long been characterized by extensive earmarking, or assignment of particular revenues to particular expenditures. Such earmarking is particularly important at the municipal and departmental levels. McCleary and Uribe Tobon (1990), for example, estimated that two-thirds of municipal and four-fifths of departmental current income were earmarked (if net public enterprise profits were included). Much of this subnational earmarking reflects central, not local, decisions. National laws require most departmental own-source revenues and a substantial fraction of municipal own-source revenues to be spent on particular activities. In addition, almost all central transfers to subnational entities in Colombia are earmarked in the sense that they must be spent on specific activities in amounts (or proportions) designated by national legislation. Colombia, like other countries, offers many examples of unfunded legislative mandates, that is, expenditure obligations imposed on subnational governments with no additional funding being attached.

In operational terms, two approaches to earmarking with quite different fiscal implications may be defined (Bird, 1998). The first approach may be called *pure earmarking*, under which certain tax (or charge) revenues flow into a special fund and become the primary source of funding for a particular expenditure. There may, or may not, be a benefit rationale for thus linking revenues and expenditures. With this system, the marginal expenditure decision is, in effect, in the hands of the taxpayer-voter. Provided there is some logical (benefit) connection between the source of the funds and the services purchased with them (and there is no overriding distributive or other reason to sever this connection) such strong earmarking offers a desirable method of financing certain public sector activities by introducing the logic of the market into the budgetary process. Apart from valorization, however, there appear to be no obvious examples of

such earmarking at the subnational level in Colombia.⁵⁸

In contrast, under a second approach, which may be called *notional earmarking*, certain types of taxes (or charges) are inframarginal and are designated to help pay for particular government services, but the revenues from these taxes finance only a portion of the government service in question. Often, there is, at best, only an extremely loose connection between the growth of earmarked revenues and higher levels of government spending in the designated area. Nor is there any commitment on the government's part to forge a close link between the amount of earmarked revenues and the volume of services to which the earmarked revenues are nominally tied. Moreover, there is seldom any benefit rationale for such earmarked taxes.⁵⁹ Under such notional earmarking, the amount of earmarked revenues may rise (or fall) while the level of associated services that are provided remains the same. In contrast to the pure earmarking case described above, the marginal expenditure decision remains firmly in the hands of the relevant budgetary authorities. Most subnational earmarking in Colombia is of this type. The relevant authority in most instances remains at the central rather than the subnational level.

⁵⁸Critical to the relative success of valorization in some Colombian cities have been careful planning and execution of projects, participation of beneficiaries in both planning and managing projects, an effective collection system, and, often, significant initial financing of the valorization fund from general government revenues (so that works can be begun in a timely fashion, without requiring prospective beneficiaries to put up all the funds in advance).

⁵⁹A portion of tobacco tax revenues, for example, was earmarked to support local sports activities in Colombia. What possible complementarity is there between smoking and football, or the craving for nicotine and the need for exercise?

Whether such earmarking is effective in altering expenditure patterns depends upon whether the amount required to be spent is more or less than would otherwise be spent. The result of receiving an earmarked transfer, for instance, may be to increase spending on the designated activity by exactly the amount of the transfer, by less than the transfer, or by more than the transfer. The effects may depend upon the local elasticity of demand for the specified activity. The evidence to date in Colombia appears to be that, for the most part, the areas to which transfers are earmarked are those in which the municipalities would, in any case, spend so that the earmarking is largely unnecessary (this question can and should be studied more rigorously).

In addition, many other restrictions and mandates have been imposed on local fiscal freedom over the last few years. To mention only two examples, an anti-corruption law requires all local governments, no matter how small, to establish and maintain an 800 telephone number. The central government has also recently mandated increased payments to various local officials. The partial evidence available suggests that the overall result of these policies in at least some municipalities has been that not only all of the transfers but also some own-source revenues are spent not in accordance with local wishes but as directed by national laws (Bird and Fiszbein, 1998). Colombian experience sug-

gests three points that may be of wider applicability. First, increased transfers do not necessarily imply increased local fiscal autonomy. Local governments in Colombia now have more resources than they did before, but in at least some cases they may have less freedom to decide how much should be spent on what. Second, repeated attempts by the central government to direct local expenditure patterns (in detail are not only basically inconsistent with the stated general intent of decentralization but seem likely to ensure that decentralization (in the sense of developing effective and efficient subnational governments) will fail. Local governments will never become fiscally responsible if the central government continues to mandate their every move. Third, the concern frequently voiced in Colombia with respect to inadequate local fiscal effort is a bit puzzling given the evidence that even if local governments do succeed in raising more revenues on their own how they can spend those funds is significantly controlled by the central government. What local politician is going to be willing to alienate local citizens by levying more taxes to carry out policies on which neither politician nor citizen may place a high priority?

While the extent to which similar questions may be raised in other countries remains to be seen, it appears that care should be exercised in equating increased transfer flows to subnational governments with either increased local autonomy or —the dark side— with increased fiscal irresponsibility by subnational governments.

Annex 3

User Charges

Since many services provided by local government benefit clearly identifiable persons (or businesses, or properties) the failure to levy appropriate user charges not only results in government spending that is higher than it should be, but also in a mix of government spending that does not reflect the real needs of society as determined by citizens, rather than bureaucrats.

At least three types of local user charges, broadly defined, exist almost everywhere; namely, service fees, public prices, and specific benefit charges. *Service fees* include such items as license fees (marriage, business, dog, vehicle) and various small charges levied by local governments for performing specific services (registering this or providing a copy of that) for identifiable individuals (or businesses). In effect, such fees constitute cost reimbursement from the private to the public sector. Charging people for something they are required by law to do may not always be sensible (for example, if the benefit of, say, registration is general and the cost is specific) but, on the whole, there is seldom much harm, or much revenue, in thus recovering the cost of providing the service in question.

In contrast, *public prices* refer to the revenues received by local governments from the sale of private goods and services such as water (other than the cost-reimbursement just described). All sales of locally provided services to identifiable private agents (from public utility charges to admission charges to recreation facilities) fall under this general heading. In principle, such prices should be set at the competitive private level, with no tax or subsidy element included unless doing so is the most efficient way of achieving public policy goals, and even then it is best if the tax subsidy element is accounted for separately.

A third category of charge revenue may be called *specific benefit taxes*. These revenues are distinct

from service fees and public prices because they do not arise from the provision or sale of a specific good or service to an identifiable private agent. Unlike prices, which are voluntarily paid (although like fees paid for services that may be required by law) taxes represent compulsory contributions to local revenues. Nonetheless, specific benefit taxes are (at least in theory) related in some way to benefits received by the taxpayer, in contrast to such general benefit taxes as those on fuel levied on road users as a class or local general business or property taxes viewed as a price paid for local collective goods. Examples abound in local finance: special assessments land value increment taxes, improvement taxes, front footage levies, supplementary property taxes related to the provision of sewers or street lighting, development exaction and charges, delineation levies, and so on. Most such charges are imposed either on the assessed value of real property or on some characteristic of that property (its area, its frontage, its location).

Unfortunately, it can be surprisingly difficult to set such prices correctly, and most fees and charges now imposed in most countries of the region are clearly defective.⁶⁰ Water rates, for example, are frequently applied as fixed charges independent of the volume of water consumed. Since the marginal cost of consumption is then zero, the result is likely to be over consumption of water and over investment in water capacity. Even when water consumption is metered, prices will usually be less than marginal cost for large water consumers if declining block rates are used. The fact that sewer charges are usually pro-rated on the amount of the water bill only compounds this pricing error.

⁶⁰Bird (1999b) provides an example with respect to San Juan Province in Argentina. For more detailed discussion of local user charges, see Bird and Tsiopoulos (1997).

Correct pricing can be quite complex. Generally, user fees should correspond to the marginal costs of providing service to each consumer. The “postage stamp pricing” approach (uniform everywhere) almost never makes economic sense, although it may sometimes be warranted when administration and enforcement costs are taken into account. Distance from the source of supply, for example, should matter in setting an appropriate user fee, as should the time of year if there are seasonal peaks in demand. Local governments should be encouraged to adopt an appropriate fee-for-services approach wherever possible. Many municipal waste management facilities and parking lots, for example, are seriously under priced. To cover capital or fixed costs, a connection (or admission) fee should be charged. Such a two-part pricing policy is usually a more efficient pricing instrument than a user fee set at the level of average total cost, but it is by no means simple to set such prices correctly.

In most countries, public sector prices are far too low, and indeed often zero. Less use is made of

charges than seems warranted because user charges are often thought to produce adverse distributional effects. Namely, if everyone must pay the same amount for a service that was previously provided free of charge, then low-income households are hurt disproportionately. Nonetheless, providing the wrong public services inefficiently is no way to improve income distribution. Moreover, it is often far from clear that benefits from a “zero-price” policy. Upper-income households may sometimes benefit disproportionately from the consumption of free, or low-cost, public services.⁶¹

Distributional concerns may thus sometimes be exaggerated. Nonetheless, they are real and must be explicitly considered in developing a user charge policy. Moreover, it is by no means always easy to establish precisely when and to what extent charges are appropriate on efficiency grounds or to prescribe how such prices should be set if they are appropriate. In short, subnational governments should impose user charges whenever feasible, but no one should think that to do so is either technically or politically a trivial problem.

⁶¹ Flat-rate water charges, for example, favor households with large lawns, swimming pools, and multiple vehicles to wash and not the poor. Free tuition to post-secondary educational institutions is an even more obvious example (Bird and Miller, 1989).

Annex 4

Subnational VATs

The VAT as a Regional Tax

The general sales tax now found in most countries is a value-added tax (VAT). The dominance of the VAT poses a serious problem for regional governments finance since independent subnational VATs were long considered to be either infeasible or undesirable for a variety of reasons. Some emphasized the high administrative and compliance costs of such taxes. Others stressed the possible loss of macroeconomic control and the general reluctance of central governments to share VATs. Still others emphasized the problems arising from cross-border (interstate) trade. Broadly, the argument with respect to such trade was that subnational VATs were distortionary, if levied on an origin (production) basis, and unworkable if levied on a destination (consumption) basis.⁶² Experience in Brazil has generally been taken to support this negative appraisal.

Brazil was the first country to introduce a full-fledged VAT (Guérard, 1973). Indeed, Brazil liked the idea of a VAT so much that it introduced not one VAT but several, one for the federal government (IPI) and one for each of the state governments (ICMS).⁶³ The latter were levied by

¹Interestingly, Poddar (1990) in the most thorough early discussion of this issue had demonstrated that the destination base was technically feasible, although the particular approach he discusses seems unlikely to work well in developing countries.

⁶³ IPI (*Imposto sobre Produtos Industrializados*) is levied by the federal government on industry and ICMS (*Imposto sobre operacoes relativas a Circulacao de Mercadorias e Servicos*) is levied by the states on agriculture, industry, and many services. In addition, another tax, the ISS (*Imposto Sobre Servicos*) is levied on a gross receipts basis by municipalities on a variety of industrial, commercial, and professional services. There are also a number of other taxes on financial transactions, retail sales of fuel and so on that are not discussed here.

producing states, so the tax was on an origin basis. Even so, complex technical and administrative problems arose as to how to apply different VATs in different states (although all are levied at a uniform rate) in addition to a federal VAT. Over the years, these problems were partly resolved in various ways, but the resulting patchwork of VATs became increasingly unsatisfactory. As a result, sales tax reform is again high on the fiscal agenda in Brazil.

The only functioning destination-base subnational VAT now in existence is in the province of Quebec in Canada (Bird and Gendron, 1998). The Quebec Sales Tax (QST) and the federal VAT (Goods and Services Tax, or GST) as they now exist in Canada constitute an operational “dual VAT” system with none of the problems usually associated with such systems. The rates of the two taxes are set independently by the respective governments. The tax bases are also determined independently, although they are essentially the same. From the beginning, both taxes have been collected by a single administration (the Quebec Department of Revenue).⁶⁴ Taxes on interprovincial sales from one business to another are basically handled by a deferred-payment system very similar to that now applied in the European Union. Exports from Quebec, whether to another province or another country, are zero-rated. Imports into the province from other provinces, or from abroad, are taxable, but the tax is assessed on interprovincial imports only when there is a sale by a registered trader to an unregistered trader or consumer in the province. Although special

⁶⁴The federal share is turned over to the federal government after deducting an agreed (per registrant) administrative cost. No problems have arisen from such intergovernmental collection arrangements, which are quite common in Canada (although the revenue flow has invariably been the other way in the past, that is, from the federal to the provincial governments.)

regimes apply to automobiles and a few other cases (Canada, 1996), in general no attempt is made to collect taxes on interprovincial purchases made directly by final consumers.

What makes this system work, Bird and Gendron (1998) argue, is the existence of the overriding federal GST as an enforcement mechanism. Audit priorities for GST are established by the federal government, but a final audit plan is agreed with the Quebec government, with the latter actually carrying out the audit and reporting the results to Revenue Canada. Since the QST is applied to a GST-inclusive base, Quebec has some direct incentive to monitor the GST as well as the QST. On the other hand, although Quebec cannot directly monitor the other end of interprovincial sales, the normal process of GST audit (carried out interprovincially by Revenue Canada) serves as a cross-check to ensure that QST has not been evaded. In effect, the existence of a federal sales tax on a more or less uniform base provides some control over interjurisdictional sales for purposes of both provincial and federal taxes. Reportedly, the system is working quite well.

Canada has thus demonstrated that with good tax administration it is feasible to operate a VAT at the subnational level on a destination basis, at least for large regional governments. A single central administration and a common base would probably be most efficient, but this degree of convergence is not essential. What *is* critical is either a unified audit or a very high level of information exchange. Most importantly, from the perspective of improving accountability, each taxing government should be able independently to determine its own VAT rate. But what can be done when there is no realistic prospect of “good” tax administration, and especially not at the subnational level? As Varsano (1999) demonstrates, a promising approach is to impose what is, in effect, a supplemental central VAT, which McLure (1999b) labels the compensating VAT or CVAT. This proposal has the major virtue of protecting the revenue when tax administration (at all levels of government) is far from well developed. Specifically, it reduces the risk that households (and unregistered traders) in any state can dodge state

VAT by pretending to be registered traders located in other states.

How might such a CVAT work? Briefly, states would zero-rate international interstate sales, but the latter would be subject to the central CVAT (as well as the central VAT). Domestic sales would thus be subject to central VAT and either state VAT or central CVAT. There would be no need for any state to deal explicitly with any other state nor, generally, would there be any need for interstate clearing of tax credits.⁶⁵ Registered purchasers in the other state would be able to credit CVAT against central VAT. The results of this procedure are twofold. First, the central government, which first levies CVAT and then credits it, would gain no net revenue from it.⁶⁶ Second, the state VAT applied to resale by the purchaser would be that of the destination state. In other words, the results are exactly the same as in the GST/QST case described above (a destination subnational VAT is applied) but the CVAT now acts to protect state revenues from some obvious frauds.

This simple system seems to make subnational VATs quite feasible and potentially attractive in developing countries (at least in large countries in which states have major expenditure roles, the VAT is the principal source of actual and potential revenue, and tax administration is not up to Canadian standards). If subnational governments are relatively homogeneous, however, or if the country is not very large, a much simpler approach would appear to be that followed in some of Canada's smaller provinces, in which a uniform subnational VAT is imposed and collected together with the central VAT and the revenues shared among subnational jurisdictions in accordance with some agreed formula. In essence, the result is similar to that in countries such as Germany and

⁶⁵This assumes that the state VAT rates are lower than the central rate. If, as in Brazil, the state rates are substantially higher, there might be some residual need for a “clearing house” (though on an aggregate, not transaction basis), but this would not seem to be a very difficult problem if, as would seem generally advisable, there is central administration of state VATs .

⁶⁶Presumably, as in Canada, the central government would receive an agreed fee for its services.

Morocco, in which a share of central VAT revenues is set-aside as a “pool” to be divided among subnational governments by formula.

Taxing Local Business

Although economists are often skeptical of the economic case for local business taxes, the political realities of governing in a democratic society are such that virtually all subnational governments wish to impose such taxes. Subnational business taxes are not only widespread, they are also generally popular with officials and citizens alike, for at least two reasons. First, they often produce substantial revenue and, in particular, tend to be much more elastic than, for example, property taxes. Second, since no one is quite sure of the incidence of such taxes, it is easy to assume, or assert, that they are paid by someone other than local residents.

The economic case for local business taxation is simply as a form of generalized benefit tax. Such benefit taxes are not only allowable at the subnational level but are essential to the attainment of efficiency (Oates, 1998). Where possible, specific public services benefiting specific business enterprises should be paid for by appropriate user charges. But where it is not feasible to recoup the marginal cost of cost-reducing public sector outlays through user charges, some form of broad-based general levy on business activity may well be warranted. It is difficult, however, to find any support along these lines for taxing any one input, whether labor (payroll tax) or capital (corporate income tax). Instead, this line of reasoning suggests that a broad-based levy neutral to factor mix such as a tax on value-added, should be imposed.⁶⁷ The most appropriate form of VAT for this purpose would be a “value-added income tax” or a VAT levied on the basis of income (production, origin) rather than consumption (destination). Compared to a conventional VAT, such a tax (which shall henceforth be referred to as a “business value tax” or BVT to reduce acronymal confusion) has three important distinguishing

⁶⁷Interestingly, the original conception of the VAT (by Adams, 1918) was as a business benefit tax. For further historical references, see Bird (1999a).

features. First, it is levied on income, not consumption: that is, it is imposed on the sum of profits and wages, or to put it another way, on investment as well as on consumption.⁶⁸ Second, it is imposed on production, not consumption: that is, it is imposed on an origin not destination basis and hence, in effect, taxes exports and not imports. And third, it is assessed by the subtraction (or addition) method on the basis of annual accounts rather than on a transaction or invoice-credit method.

While the danger of tax exporting and, more importantly, beggar-my-neighbor tax competition suggests that it may be advisable to place a floor (and perhaps also a ceiling) on such taxes, this form of business taxation is likely to be less distorting than such existing subnational business taxes.⁶⁹ As a replacement for such taxes, a BVT would improve subnational tax systems in several ways. First, it would be more neutral than most existing local business taxes, which often discriminate against (certain forms of) investment. Second, it would likely be less susceptible to base erosion. Third, a BVT should be more sensitive to cyclical realities than most other forms of business tax (e.g. capital taxes, property taxes). All these advantages would apply even if the rate of the BVT were set to produce the same revenue as existing business taxes. If the rate were set to match roughly the benefits received basis as suggested above, it would generally be lower, and

⁶⁸Note that an income basis can be applied to gross income, net (of depreciation) income, or even consumption (add back depreciation and deduct capital expenditures). It is assumed here that the net income basis is the most logical.

⁶⁹These arguments are by no means solely theoretical, since there is already some important real world experience with such taxes, both in the states of Michigan (the former BAT, or business activities tax, from 1953 to 1968, and, since 1976, the SBT, or single business tax) and New Hampshire (the business enterprise tax, or BET) as well as in Italy (the new regional business tax, the IRAP) and Germany (the old local trade tax or *gewerbesteuer*). For further discussion, see Bird (1999a).

the tax would have the additional important advantages of eliminating inefficient spillovers and

encouraging more responsible and accountable subnational governments (Bird, 1999a).

Annex 5

Equalization Transfers in Canada and Australia

Revenue Equalization in Canada

The Canadian Constitution states that “Parliament and the Government of Canada are committed to the principle of making equalization payments to ensure that provincial governments have sufficient revenues to provide reasonably comparable levels of public services at reasonably comparable levels of taxation.” Although this provision was added to the Constitution in 1982, an equalization program had actually been in place since 1957, when it was introduced as part of a realignment of taxing powers following the extreme centralization that had evolved during the war years.

In 1967, Canada's equalization system took what is essentially its present form with the adoption of the so-called “representative tax system” (RTS) approach (an approach originally developed by the U.S. Advisory Commission on Intergovernmental Relations but not, in fact, used in that country). Under the RTS approach, the fiscal capacity of each province is calculated each year (by the federal Department of Finance) as the difference between the sum of its per capita fiscal capacity for each separate revenue source and a standard capacity measure based on the average per capita fiscal capacity of the “middle” five provinces (excluding oil-rich Alberta and the four small and poor eastern provinces). The system works as follows:

(1) Define and quantify all revenues with respect to which capacity is to be measured. At present there are over 30 separate revenue sources for which this calculation is made.

(2) Define a tax base for each revenue source, which reflects the tax base, typically used in practice. In some instances, as with the personal income tax, where the base is essentially uniform across the country, this task is simple. In others, as with respect to property taxes, it may be very dif-

ficult to define and calculate a credibly comparable tax base.

(3) Calculate a notional “national average tax rate” for each revenue source as the ratio of actual taxes collected by all provinces for the total standardized tax base.

(4) Apply this average tax rate for each source to the standardized base for each province to estimate that province's capacity for that source.

(5) Add up all the separate calculations to derive an estimate of total revenue capacity for each province.

(6) Divide this total estimate by the population of the province to obtain per capita revenue capacity.

(7) Compare these capacity estimates with the five-province standard mentioned above. If a province's per capita capacity is below the standard, it receives an equalization payment sufficient to bring it up to that standard. On average, this means increasing its capacity to over 90 percent of the per capita national average capacity.

Four aspects of this equalization system deserve emphasis. First, no attempt is made to measure differences in expenditure needs. This has been considered several times but has been rejected largely on the grounds that the data required to measure such differences are not very reliable. Second, although provinces below “standard” capacity receive equalization transfers, provinces with above-standard capacity are not penalized: equalization is paid entirely from federal revenues. Third, in principle the total amount paid is determined by the calculations described above. In practice, however, total equalization entitlements are limited to 1.2 percent of GDP, and this limit has, on several occasions, resulted in a

scaled reduction in the equalization payments generated by the formula.

Finally, while Canada's extensive experience with this system is generally considered a success and seems to be well accepted by most relevant actors in the political process, substantial technical criticisms have been made of various aspects of the system and it is not impossible that various changes may be made in the future (for example, to incorporate expenditure needs, to move from the RTS to some other basis of capacity calculation to include other federal transfers in the sources for which equalization is calculated, or to reduce the extent to which the system biases provincial tax policy choices).

Expenditure Equalization in Australia

The most important difference between the Canadian and Australian equalization systems is that Australia, unlike Canada, attempts also to equalize for differences in expenditure needs. The procedure is more or less the same as that outlined above with respect to revenue equalization. Since 1981, the Grants Commission (see Annex 8) has assessed every five years the needs of all states relative to "standard" (average) per capita recurrent expenditures. While, as in Canada, the precise list of items equalized changes from time to time, the number of expenditure items equalized tends to be longer than the number of reve-

nue items. In the original calculation of these "grant relativities," for example, 58 expenditure items were assessed compared to 23 revenue items (Bird, 1986). Subsequently, the number of items included in both categories was reduced, and the methodology altered slightly, but the procedure remained largely unchanged, in principle.

The expenditure "need" for a particular service is calculated for each state by calculating the differential cost for that particular state of providing services of a standard level, range, and quality. The per capita differential cost is the difference between per capita "standardized" expenditure (the amount the state would need to spend, if operating at average efficiency, to provide a standard level and range of the service) and per capita "standard" expenditures (the average expenditure of all states weighted by population). Among the "need" factors taken into account are scale, population dispersion, urbanization, age structure, percentage of aboriginal population, and environmental factors. To some extent, these calculations are inevitably judgmental or subjective in nature. Moreover, other technical and philosophical criticisms have frequently been made of this system, not least that its complexity makes it virtually impossible to understand and inappropriately turns over important resource allocation decisions to a nonelected body. Nonetheless, Australia's long and generally successful experience with this system demonstrates that expenditure equalization is feasible.

Annex 6

Capitation Grants for Education

Assume that the provision of a certain basic level of primary and secondary education to all citizens is a fundamental goal of the central government, but that this function is, for whatever reason, to be delivered by subnational governments. How can transfers be used to help achieve this goal? One possibility is to establish a per capita transfer based solely on the cost of providing the desired standard of service to the relevant population. In a system along these lines proposed recently in South Africa, for example, the approach taken was to establish a minimum amount per student that should be spent by each province on education and to ensure that the required funding was made available through a national transfer. The minimum standard was determined by estimating the average cost of teachers at the national (not provincial) average salary required to permit all school-age children (excluding those in private schools) to attend school, assuming a 35:1 ratio of pupils to teachers in secondary schools, and 40:1 in primary schools. The grant thus estimated was then increased by 29% to allow for costs other than teachers' salaries. Provinces were to be free to spend more than this if they wished to do so out of other resources, but they were expected to spend at least this amount. A somewhat similar scheme was proposed a few years ago in Colombia (World Bank, 1996a), and a system along these lines already exists in Chile.

The factors that need to be determined in designing such a transfer system include the unit cost of provision, the number of units provided, the rate of extension of coverage, adjustments to unit cost over time, the division of transfer between different subnational levels of government, and the rate of transition. A brief discussion of these factors illustrates some of the critical detail that must be taken into account in designing an actual transfer system.

Unit Cost. One way to determine this cost might be simply to divide the existing recurrent costs in education by the number of students currently served. Alternatively, this cost might be determined, as mentioned above, by assuming a target teacher-pupil ratio, an average teacher cost, and an expansion factor per teacher for non-salary costs. In principle, it would likely be best to have a standard national cost. It may be possible to accommodate a limited amount of variation (regional, rurality, density) without unduly complicating the formula. On the other hand, a uniform national minimum guarantee might avoid any appearance of regional bias. For example, the central government could deliver throughout the country exactly the same transfer per unit (child) served.

Number of Units. "Unit" refers to the quantity whose cost was calculated in the preceding exercise. The unit might be the number of students currently enrolled in official schools. The number of students currently enrolled, however, may be a poor indicator of the extent to which the center should be financing education in a particular municipality. A better indicator of need might be the number of persons of school age, although clearly care would have to be taken to ensure that the jurisdiction has adequate incentives to actually provide schooling for these children.⁷⁰ Some adjustment may also be needed to take into account students accommodated in the non-public school system.

⁷⁰One way to do this might be to provide the transfer directly to the potential students conditional on their actually attending an appropriate educational institution, but to pay it to the institution if the student is actually enrolled, and attending. This is a form of "voucher" system. Something along these lines was recently proposed for the health system in Colombia (World Bank, 1996a).

Extension of Coverage. In most countries, different territorial units at present have different degrees of coverage. A goal of national policy might well be to extend the coverage of the basic education system to all eligible students. The question is how best to accomplish this goal within the framework of a per capita grant in support of education. Should the grant be paid from the beginning in terms of potential students? Or should it be paid initially in terms of actual enrolments with an adjustment for the extension of coverage? If the latter approach is taken, over what time period should the transition to full coverage take place? And how should the transitional increment be allocated between territorial units — for example, proportional to the size of the “gap”?

Adjustment of Unit Cost. Should the initially determined unit cost be adjusted over time? Presumably, at a minimum it should be held constant in real terms. Assuming the expansion of

coverage has already been taken into account, there would seem no reason to provide for any systematic adjustment procedure other than indexation,

Transition. More policies probably fail because of inadequate attention to transitional issues than because of bad design. At least three transitional dimensions need to be taken into account: (1) the pace at which any service “deficits” (below-average levels of provision) are rectified; (2) the pace at which “excesses” are corrected; and (3) the resolution of the imbalance if (1) does not equal (2). As a rule, it is easier to build up than to reduce services.⁷¹ Since reducing government expenditures is never politically easy, it may be necessary to reduce transfers (if this proves to be one implication for some regions of a more equitable system) more slowly than they are increased, which may require some additional funding for a transitional period.

⁷¹Of course, it is also important not to dump too much new money too quickly into any governmental unit, since it will always be spent, if not necessarily efficiently.

Annex 7

Matching Grants and Infrastructure Finance

Central governments have two distinct reasons for being interested in what local governments do with respect to infrastructure. First, some local infrastructure projects may involve significant externalities. Second, some projects may constitute essential elements of national development programs. Infrastructure related to the provision of basic education and health services, for example, may qualify for both reasons, as may projects improving the level and quality of water supply and sewerage treatment. Support of local roads and some rural development projects may be justified as part of efforts to improve the economic productivity of poor rural areas. In a decentralized system, subnational governments should identify infrastructure needs and execute projects. Financing large infrastructure projects from local resources alone may not be possible, given the scanty current revenues of most local governments. Moreover, small localities seldom have much access to private capital markets. If they are to carry out costly public works, as a rule local governments depend heavily upon grants from higher-level governments.⁷²

Local government boundaries seldom coincide with the benefit areas of their activities. One way to deal with interlocality benefit spillovers is through the use of matching grants.⁷³ Indeed, the

⁷²Subsidized loans combine the elements of a matching grant with a commercial loan. Although in many countries local capital finance through borrowing takes place mainly from government-sponsored and financed agencies, the record of such agencies is poor, with many loans not being repaid and local governments having few incentives to pay (Davey, 1990). On the whole, it seems better to keep any such borrowing close to commercial terms and to leave any subsidy to the grant side.

⁷³Another is by boundary changes for example, the creation of special districts to provide particular services. While this approach is sometimes useful, it should be remembered that such jurisdictions may

transfer rationale with the strongest basis in the economic literature is that the local activity in question may spill over to other jurisdictions.

Under this approach the correct “matching rate” (the degree of central subsidy) is set by the size of the spillovers. A matching grant program designed to encourage the optimal provision of public services, since it depends upon the level of associated externalities, should vary primarily with the nature of the activity. However, no country has achieved full equalization of local fiscal capacities, therefore, a uniform matching level offering the same “price” to all local governments will yield non-uniform responses in rich and poor localities (Feldstein, 1975). Indeed, even if local revenue bases are fully equalized, need or cost differentials may call for some equalization element in matching grant formulas. For example, per capita grants for roads in sparsely populated and mountainous regions may be larger (as in Switzerland, for example) because the per capita cost of achieving any particular standard of road service will obviously be higher. Such modifications to ensure the desired allocative results from matching formulas should not be confused with redistributive transfers.

A more basic problem with matching grants as a way of dealing with spillovers lies not in such concerns but in the simple fact that no one has a

undesirably weaken general-purpose local governments both in terms of competition for resources and reduced political accountability, and hence hamper the efficient provision of other public services. A recent examination of Canadian experience with special-purpose districts concluded that they made government more difficult for citizens to understand; they reduced the degree of control citizens have over government; they reduced accountability and hence, the overall efficiency of resource allocation in the public sector; and they appropriated significant revenue that would otherwise accrue to local governments (Kitchen, 1993).

good idea of the magnitude of spillovers associated with particular services. In practice, the spillover argument at most provides a rationale for central government support of some local expenditure. It almost never indicates *how much* support is needed. Nonetheless, the required matching ratios found in many countries (say, 10% or 20%) as a rule seem likely to be far too low for allocative efficiency.

Most countries could improve the assistance provided through matching transfers for purposes of capital expenditure in several ways: First, the terms and conditions of such transfers should require local governments to prepare both an adequate investment plan and an adequate maintenance plan (as well as an appropriate user charge policy). Second, localities that receive such transfers should be selected not by political factors but by an objective process that pays attention to need and capacity factors as well as to the economic evaluation (such as cost-benefit analysis) of the project in question. Third, adequate technical assistance (not necessarily, and indeed, not even desirably from central governments) should be made available to local governments to permit them to develop plans, arrange financing, manage construction, and operate the facility in the most efficient fashion possible. Fourth, the execution and operation of grant-aided work should be monitored and evaluated, with periodic progress reports, field inspections, and formal evaluations of outcomes. Finally, all local governments receiving such aid should be required to provide surveys of the condition of the infrastructure to which aid is to be directed in order to permit adequate assessment of future needs.

Local governments, even when (as is often the case) they have not been involved in the selection or execution of projects, are sometimes assumed by the central government to be willing and able to look after the subsequent costs required to keep the infrastructure operating and in good condition. This assumption is often mistaken. Local governments may not have the financial resources or technical capacity to undertake this task, and the incentives facing them seldom encourage them to do so. Perhaps the worst of all systems is that in which the infrastructure is provided at no cost to the locality whose only contribution is financing recurrent operation and maintenance costs. As noted in the text, localities are more likely to maintain facilities if they paid at least something for them and therefore feel a sense of ownership. Moreover in many countries, the political reality is that the less a local government does to maintain its infrastructure, the more likely it is to be rescued from above by the provision of yet more subsidized investment.

In some instances ingenious solutions have been found to the problem of financing the recurrent costs required to keep infrastructure in operation. Where there are well-defined projects with clear beneficiary groups who have no alternatives, solutions can often be found (see, for example, a recent study of rural road maintenance in Indonesia—UNCHS, 1996). Nonetheless, it is critical to factor in appropriate assumptions about operating and maintenance as well as cost-recovery from the beginning in the process of planning capital projects if the expected benefits of the project are to be realized. In this, as in many other ways, countries must strive to incorporate into their local government system institutions that encourage the exercise of forethought, imagination, effort and prudence if the results of decentralization are to be as beneficial in practice as they may be in theory.

Annex 8

Intergovernmental Institutions

Germany epitomizes what has been called “interlocked” federalism in which, no level of government can do anything important without securing the explicit agreement of the other level.⁷⁴ The *Bundesrat*, or second chamber of the national legislature, essentially represents the state governments, and its approval is needed for central legislation. Other important governmental functions are either explicitly or implicitly “joint” federal-state activities and depend upon the active cooperation of both levels in terms of finances and operations. Germany's administrative federalism thus deals with the “integrating” function (ensuring that the political and bureaucratic players at each level are singing from the same sheet) largely through its formal governmental structure. However, no special institutions exist to provide a comprehensible public face to German fiscal federalism, although the high level of technical expertise in academic and nongovernmental circles makes up for this lack.

Canada also has no special institutional structure to help citizens understand who does what or, indeed, what is actually going on in the complex world of Canadian public finance. While, as in Germany, this gap is partly made up by nongovernmental research bodies, the problem is more serious in Canada because, unlike Germany, it is home to different territorially-based language groups and because there are no integrating institutions at the political level. Canada has neither a formal institutional structure such as the *Bundesrat* nor a less formal but equally effective political structure (such as the Swiss party system) to keep

federal and provincial policies in tune with each other, let alone to ensure that each side is operating in full knowledge of what the other is doing, or intends to do. One result is that virtually every issue in Canadian politics gives rise to federal-provincial conflict.

At the bureaucratic level, which is in many ways what matters most for the equitable and efficient delivery of services, Canada's executive federalism works extremely well, owing in part to the proliferation of intergovernmental committees and agencies at the operative level. With respect to fiscal planning, for example, there are regular meetings of federal and provincial officials working within a common framework.

With the recent disappearance of the Advisory Commission on Intergovernmental Relations (ACIR), the *United States* too now has no publicly-supported body focusing chiefly on federal-state-local issues. The ACIR never played a significant role in the formation or monitoring of U.S. fiscal performance, although it did provide numerous useful studies of intergovernmental fiscal relations. The Commission had rotating memberships from Congress, states, and local governments, making it difficult to speak with one consistent voice on policy issues. Sometimes big cities were represented on the Commission, sometimes counties, and sometimes states. Further, the staff's expertise was mainly in public administration, not public finance or public accounting, so that most analyses focused on the administrative structures of government. Finally, the published studies tended to be one-time reports without follow-up.⁷⁵ Their impact on the policy process was minimal.

⁷⁴ What is said of Germany is, on the whole, even more true of Austria. In contrast, Switzerland achieves similar “consensual” results in a totally different institutional fashion, partly owing to the nature of its political party system and its “committee” presidency and partly through its very detailed constitution (Bird, 1986).

⁷⁵ For an example that is particularly interesting in the present context, see the study of “ACIR” type agencies at the state level in ACIR (1981).

As in Canada, one problem of American federalism in recent years has been an undesirably intrusive federal government. Although the Courts have been seen as the primary protector of state and local government policy interests in the United States, the judiciary has never articulated any firm principle of governance to guide decision-making. In reality, the federal Congress remains the dominant factor in fiscal policy making, with congressional decisions being largely decided by national political needs—whether for budget control (as in the recent decentralization of welfare spending) or ensuring each representative a favored program for local constituents (as in recent highway legislation). An interesting recent proposal suggested that a “Federalism Impact Statement” should be attached to all such federal initiatives in order to reduce unnecessary federal intrusion on state spheres of action (Inman and Rubinfeld, 1997).

Probably the best-known intergovernmental institution is the Commonwealth Grants Commission (CGC) in *Australia*. A balanced appraisal might be that the CGC has served a useful function in depoliticizing, to a limited extent, some of the discussion about the distribution of federal grants (in part, unfortunately, by making the process so complex and academic that it sometimes seems that no one can quite understand it—Mathews and Grewal, 1997). It has also done useful work in analyzing and suggesting remedies for such specific problems as financing the Northern Territory and the Capital District and in compiling detailed comparative public finance information. The prestige that the CGC has developed in these matters over the years is such that their conclusions have almost always been acted upon by the federal government, although in recent years they have not been received with favor by some of the richer states (New South Wales and Victoria). For some years, Australia also had an Advisory Commission on Intergovernmental Relations which carried out some broader studies on federal matters, as did a federally-funded Federalism Research Centre (FRC) at the Australian National University, but both these bodies have now fallen by the wayside. Nonetheless, as in Canada, Germany, and the U.S., and in part perhaps because of the prior support for federalism

research through such bodies as the FRC, Australia has a significant cadre of nongovernment researchers who understand federal finance thoroughly and contribute substantially to public debate on these issues.

In principle, another important intergovernmental agency in Australia is the Australian Loan Council (ALC). In fact, however, the ALC appears largely to act as an instrument through which the federal government controls subnational borrowing. Originally established in the 1930s to coordinate state and federal economic and financial policies, the ALC's role in this respect diminished sharply when the federal government took full control of fiscal matters during the Second World War. Federal control was loosened somewhat in the late 1970s for state and large local governments and further deregulation in 1984 removed borrowing limits and extended subnational access to both domestic and foreign markets. Nonetheless, federal control still exists in the form of global borrowing limits (Loan Council Allocations) established for each state. These limits, which are based not only on capital funding needs and comparative fiscal conditions but also on macroeconomic considerations, are effectively set by the federal government.

Turning to developing countries, *India* provides the best-known example of a not uncommon device (Nigeria has a similar provision) of providing for a periodic review of federal financial matters by a specially appointed finance commission (FC). India has now had 10 such commissions, one every five years, and the 11th is scheduled to report shortly. Sometimes the reports of these commissions have been informative and useful to those interested in federal finance, sometimes they have been followed by the central government, and sometimes both or neither of these things has been true. As in the case of the CGC in Australia, however, the scope of the Finance Commission has been limited to a narrow subset of federal-state fiscal arrangements. Some observers have argued that even the best of their reports have generally been more than offset by divergent policies followed by other bodies, most notably the Planning Commission (Rao and Sen, 1996). Moreover, in India, as in Australia, the commis-

sions have not really played any role in integrating state and federal policies to any extent. In principle, *Pakistan's* Finance Commission should do better in this respect since it incorporates federal and provincial legislators. Although political circumstances in Pakistan have often made the Commission irrelevant, when the FC has operated fairly well its recommendations have

generally been notably simpler and more transparent than those produced by the more academic, less political, exercises in India and Australia. Finally, *South Africa's* new Constitution set up a Financial and Fiscal Commission (FFC). Initially, the FFC usefully set out some of the key issues in creating a new decentralized and democratic state. As yet, however, it has had relatively little impact on either public reality or political action, in part perhaps to a considerable extent the central government has never really accepted the provinces as having any real role to play.