Decentralization and Infrastructure in Developing Countries: Reconciling Principles and Practice

Roy Bahl and Richard Bird
Andrew Young School of Policy Studies, Georgia State University, and Institute on Municipal Finance and Governance
The Institute on Municipal Finance and Governance (IMFG) at the Munk School of Global Affairs at the University of Toronto focuses on developing solutions to the fiscal and governance problems facing large cities and city-regions. IMFG conducts original research on Canadian cities and other cities around the world; promotes high-level discussion among Canada’s government, academic, corporate, and community leaders through conferences and roundtables; and supports graduate and post-graduate students to build Canada’s cadre of municipal finance and governance experts. It is the only institute in Canada that focuses solely on municipal finance issues and large cities and city-regions. IMFG is funded by the Province of Ontario, the City of Toronto, Avana Capital Corporation, and TD Bank.

The IMFG Papers on Municipal Finance and Governance are designed to disseminate research that is being undertaken in academic circles in Canada and abroad on municipal finance and governance issues. The series, which includes papers by local as well as international scholars, is intended to inform the debate on important issues in large cities and city-regions.

Roy Bahl is Regents Professor of Economics, emeritus, Andrew Young School of Policy Studies, Georgia State University. His research interest is the public finances of developing countries.

Richard Bird is professor emeritus in the Rotman School of Management and Senior Fellow of the Institute on Municipal Finance and Governance at the Munk School of Global Affairs, University of Toronto. His research interests include tax policy and administration, local and regional finance, and intergovernmental fiscal relations.

Acknowledgement
An earlier version of this paper was prepared for the Dana Weist Memorial Conference on “Decentralization and Infrastructure: From Gaps to Solutions,” sponsored by Georgia State University and World Bank, February 13–14, 2013, Washington, D.C. The present version has benefited from helpful comments from the conference organizers, Jonas Frank and Jorge Martinez-Vazquez, as well as from our conference discussants, Berndt Spahn and François Vaillancourt, and others at the conference in addition to the careful editing of Philippa Campsie.

The views expressed in this paper are those of the authors and not those of the conference sponsors, IMFG, or its funders.
Papers on Municipal Finance and Governance


Decentralization and Infrastructure in Developing Countries: Reconciling Principles and Practice

Roy Bahl and Richard Bird

Abstract
The paper first sets out the theory of expenditure assignment with respect to decentralizing infrastructure expenditure, and then considers how practice around the developing world appears to differ from what that theory appears to suggest. We suggest several ways in which theory and practice might be brought closer together. The most important is simply to begin by taking a more comprehensive approach to infrastructure reform in countries in which much such investment is, properly, at the subnational level. Two critical preconditions for effectively decentralizing investment are, first, clear assignment of infrastructure responsibilities and, second, effective local government accountability. Neither condition is now satisfied in many developing countries. One key reason is that few countries have developed appropriate and adequate local government revenue systems. Although much attention has been paid to financing infrastructure through borrowing and public-private partnerships (PPP), and such approaches may have important roles to play in developing adequate infrastructure in some countries, they can neither substitute for a sound local revenue system nor realize their full potential in the absence of such a system. In addition, since local governments are seldom equal in fiscal or economic terms as a rule, effectively decentralizing infrastructure requires the careful development of different (asymmetric) approaches for different sizes and types of subnational government structures.

Keywords: decentralization, expenditure assignment, infrastructure investment, borrowing, public-private partnerships

JEL codes: H72, H77, O21, O23, R53
Decentralization and Infrastructure in Developing Countries: Reconciling Principles and Practice

1. Introduction
This paper reviews the theoretical and practical issues surrounding the decentralization of responsibility for infrastructure to local governments, with particular emphasis on local governments in developing countries. The focus is on structural issues rather than day-to-day management issues: the latter are important but even the best management is unlikely to produce the best possible results if the structure of local governance and finance is flawed. We begin by setting out the theory of expenditure assignment with respect to decentralizing infrastructure expenditure, and then consider how practice around the developing world appears to differ from what that theory appears to suggest. In the balance of the paper, we suggest several ways in which theory and practice might be brought closer together. The most important is simply to begin by taking a more comprehensive approach to infrastructure reform in countries in which much of such investment is, properly, at the subnational level. Two critical preconditions for effectively decentralizing investment are, first, clear assignment of infrastructure responsibilities and, second, effective local government accountability. Neither condition is now satisfied in many developing countries. One key reason is that few countries have developed appropriate and adequate local government revenue systems. Although much attention has been paid to financing infrastructure through borrowing and public-private partnerships (PPP), and such approaches may have important roles to play in developing adequate infrastructure in some countries, they can neither substitute for a sound local revenue system nor realize their full potential in the absence of such a system. In addition, since local governments are seldom equal in fiscal or economic terms as a rule, effectively decentralizing infrastructure requires the careful development of different (asymmetric) approaches for different sizes and types of subnational government structures.

According to the standard economic theory of expenditure assignment, under the right circumstances, certain services and the infrastructure needed to provide them can be more efficiently provided by subnational governments than national governments. In industrial countries, infrastructure investment is indeed decentralized to a significant extent (Estache and Sinha 1995). In OECD countries, even big projects, such as ports and airports, are often locally managed and funded (Bel and Fageda 2009). In the European Union, for example, the share of subnational investment in economic infrastructure is 60 to 70 percent in the older member states and 40 percent in the new member states; subnational shares in social investments (e.g., in schools and hospitals) are even higher (Kappeler et al. 2012).

In low- and middle-income countries, however, economic theory provides less guidance for best practices. This is mostly because the assumptions underpinning
the decentralization model do not fit the developing country setting as well. The result is a great deal of variation in the extent to which responsibility for infrastructure service delivery is assigned to local governments (Alm 2010). In the countries listed in Table 1, for instance, the equivalent of a large share of subnational own revenue is devoted to infrastructure. In fact, the amount spent on infrastructure by subnational governments in Peru is greater than the amount they raise from own sources.

Even in those developing countries in which subnational governments are important investors in infrastructure, it is by no means obvious whether expenditures on infrastructure have been over-assigned or under-assigned to such governments, or whether the right expenditures have been devolved. The open question is whether a change in governance arrangements (or in the incentives to policy-makers that lie beneath these governance arrangements) would lead to an improvement in the delivery of infrastructure services. The institutional arrangements in question include not only the structure of subnational governments, but also the nature of political representation, spending and revenue mobilization powers, borrowing restrictions and practices, the degree of autonomy in management, the design and implementation of intergovernmental transfer systems, and the accountability of governments at all levels to those for whom they presumably speak.

1. In the aftermath of the financial crisis, many questions have been raised about the relationship between decentralization and public investment in a time of increasing budgetary pressure. A useful recent survey of this literature may be found in Fernandez Llera (2012), but we do not discuss the issue here.

2. For example, Afonso, Araujo, and Junior (2005) argue with respect to Brazil that “the shortage of investment in infrastructure is even more serious when we consider that an increasing, and already the major, part of expenditure on capital formation by public sector authorities has become decentralized… Institutionally, these regional governments do not have the competence to concede, regulate, or carry out functions in the majority of actions and services that are classified as infrastructure related (with the exception of sanitation).”

<table>
<thead>
<tr>
<th>Country</th>
<th>General government investment (% GDP)</th>
<th>Subnational government investment (% GDP)</th>
<th>Subnational as share of total (%)</th>
<th>General government taxes (% GDP)</th>
<th>Subnational own revenues (% GDP)</th>
<th>Subnational as share of total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argentina</td>
<td>1.3</td>
<td>1.1</td>
<td>84.6</td>
<td>30.8</td>
<td>5.6</td>
<td>18.2</td>
</tr>
<tr>
<td>Colombia</td>
<td>1.9</td>
<td>1.3</td>
<td>68.4</td>
<td>15.9</td>
<td>2.9</td>
<td>18.2</td>
</tr>
<tr>
<td>Chile</td>
<td>1.4</td>
<td>0.2</td>
<td>14.3</td>
<td>19.6</td>
<td>1.5</td>
<td>7.6</td>
</tr>
<tr>
<td>Peru</td>
<td>2.4</td>
<td>1.2</td>
<td>50.0</td>
<td>17.4</td>
<td>0.8</td>
<td>5.0</td>
</tr>
</tbody>
</table>

Sources and notes: Investment (average acquisition of fixed capital on accrual basis; data for 1995–2006) from de Mello 2012; revenue (for 2008 only) from Gomez Sabaini and Jimenez 2012.
In Section 2 of this paper, we define what we mean by “decentralization” and “infrastructure,” discuss normative principles that guide the assignment of responsibility for delivering infrastructure services, and note how the practice observed in most countries deviates from what theory appears to suggest. Section 3 considers a number of ways in which theory and practice might be better reconciled, emphasizing in particular the need to consider the broader governance context of each country. Section 4 briefly deals with two special “asymmetrical” issues of importance in many developing countries: infrastructure provision in metropolitan and remote rural areas, and regional equity. Section 5 concludes with a summary of the policy directions suggested by this review.

2. Infrastructure and Expenditure Assignment

Though the term “infrastructure” is commonly used in the development economics literature, it is not always defined, and can take on different meanings. The term “public infrastructure investment,” for instance, is sometimes used to denote government investment, even though it also includes investment by publicly owned companies such as utilities. As a rule, however, such corporate investment is recorded as government (rather than corporate) investment only if the revenues of such firms cover less than half their production costs: that is, the source of funding, not the formal ownership, determines how investment is classified (Alegre et al. 2008). More important for the present argument, “infrastructure investment” itself is sometimes used broadly to include not only physical capital formation, but also expenditure in areas considered to increase human capital, for example, education. Even excluding such expenditure, it is often useful to think of government capital formation as falling into several distinct categories.

Alegre et al. (2008), for example, distinguish four categories in terms of the functions served by such investment: (1) Redistribution (housing, recreation, social protection), (2) Public Goods (defence, environment, order and safety, general public services), (3) Hospitals and Schools (health and education), and what they call simply (4) Infrastructure, that is traditional public works projects, of which transportation is by far the most important in quantitative terms in most countries. As they argue:

This type of government investment [public works] has the most direct economic impact by reducing firms' production and transaction costs. The economic impact of government investment in health and education sectors is more long-term and less direct in character, as it facilitates the building up and maintenance of the economy's stock of human capital. Investment in public goods affects the economy's allocative efficiency indirectly through framework conditions for productive activity. Finally, redistribution affects the economy's income distribution rather than allocative or productive efficiency per se (Alegre et al. 2008, 26-27).
Although one may quibble with some of these statements, this classification is nonetheless useful in that it distinguishes clearly between investments that have direct and indirect effects on the level and nature of economic activity. It also indicates that the most relevant government investment usually takes the form of “network” investment (e.g., transportation and communication) rather than “point” investment (e.g., hospitals, administrative buildings), even though the former is also geographically specific in its location and the latter may often also be part of a network from some perspectives (e.g., primary health centres feeding into hospitals or elementary schools into secondary schools).

For the most part, the focus in this paper is on the narrower definition of infrastructure investment as the construction, operation, and maintenance of the long-lived physical assets required to deliver such specific public services as land transportation (highways, roads, streets, bridges, and ancillary services such as street lighting, street cleaning, signage, etc.), potable water (supply, distribution), wastewater management (sewerage, disposal), and solid waste collection and disposal (including hazardous waste). Much of the discussion—for example, with respect to public-private partnerships—may also apply to other “utility” services, such as electric power generation and distribution, natural gas distribution, airports and airways, ports and navigation control systems, railroads and public transit systems, irrigation, telecommunications, and other types of investment mentioned by Alegre et al. (2008). For the most part in this paper, however, we focus on the simplest, most basic and arguably most important infrastructure investments usually assigned to decentralized governments—roads, water and sewerage, and solid waste disposal. Since infrastructure investments are, by definition, located in a particular physical location within one or more specific
political jurisdictions, even with this restricted set of activities much the same wide range of factors must be considered as with a broader definition.

The term “fiscal decentralization” also requires definition. In this paper, we interpret decentralization to mean the devolution of explicit authority and responsibility for a specific activity to a specific unit of government. Even this definition is both too broad and too narrow for our purposes, however. It is too broad in the sense that the assignment of a particular discretionary power to a level of government is almost never simply a matter of drawing lines between broad classes of activity such as “water and sewers” and “education” and assigning each exclusively to one level of government. As discussed below, each activity or function can be and usually is “unbundled” in a number of ways. At the same time, defining decentralization as simply equivalent to devolution is also too narrow because it excludes much of the important infrastructure investment undertaken by decentralized governments engaged in delivering delegated services such as health and education. Although our principal focus here is on the provision of physical infrastructure (including operations and maintenance) by subnational governments, inevitably some of the specific instances to which we refer and some of the arguments we make will stretch these boundaries to varying extents.

2.1 The basic rule of expenditure assignment: Efficiency

The basic rule of efficient expenditure assignment is to devolve each function to the lowest level of government consistent with its efficient performance. This idea is expressed in the “decentralization theorem” (Oates 1972). So long as there are local variations in tastes and costs, there are potential efficiency gains from assigning responsibility for public-sector activities to the lowest level of government possible. Under such an arrangement, local decision-makers act on behalf of voters to decide what services are provided, to whom, and in what quantity and quality. The apt phrase is that “people get what they want” under a decentralized system, with the result that overall public welfare is enhanced. A well-known example of this principle is that of “subsidiarity” in the European Union.5

For some expenditure functions, however, assignment to the lowest level of government does not lead to a welfare gain, for two main reasons: the presence of externalities and economies of scale. Although there are other ways to deal with such problems (such as intergovernmental transfers and private contracting), either might cause one to favour more centralized infrastructure spending, especially because many infrastructure projects are expensive, capital-intensive, and characterized by spillover effects.

Nonetheless, as Table 1 illustrates, in at least some developing countries, a large proportion of capital expenditures are made by subnational governments.

5. For an interesting comparison of the decentralization theorem and the subsidiarity principle, see Breton, Cassone, and Fraschini (1998).
One explanation might be the importance placed on satisfying local preferences. Another might be that fiscal centralization is not the only way to deal with economies of scale and spillover effects or other commonly cited pro-centralization factors, such as corruption or administrative capacity. Yet another explanation might be that there are overwhelming political economy considerations (e.g., regional equity concerns) for decentralizing more infrastructure investment than the conventional fiscal federalism literature might suggest.

One way around the assignment problems that arise when externalities or economies of scale are present is unbundling. This involves breaking the delivery of a function down into its subcomponents so that the most appropriate assignment can be made to local governments. For example, expenditures can be disaggregated by sectors (e.g., education, health), services (e.g., primary curative care, primary education), or activities (e.g., policy and standards, planning, asset creation, operation, and maintenance and operation). The World Bank (2007) has proposed unbundling service responsibilities for rural local governments in India by cross-classifying activities against sectors and services. The result of such an approach is to provide clarity in expenditure assignment and also to provide constituents with a clearer idea of whom to hold accountable for what service.

Boadway and Shah (2009) take a similar approach. In their schema—outlined in Table 2—only water, sewer, refuse, and fire protection are assigned to local governments, with local control over such local infrastructure (roads, water, sewage, irrigation) likely to be “more cost-effective as well as more suited to the needs of the users” (147). In addition, although their enthusiasm for public enterprises in general is restrained, Boadway and Shah (2009) recognize that local control over public enterprises providing local utility and transportation services is also appropriate, provided the users are also local (148).

Even in principle, however, matters are complicated, for two reasons. First, as already mentioned, responsibilities for some services are often shared between different levels of government (national, state, and local). As Table 2 shows, in the Boadway-Shah presentation such sharing is deemed appropriate not only with respect to police, highways, parks and recreation, education, health, and social welfare, but also even broader policy areas such as fiscal policy, regulation, and natural resources as well as—interestingly—direct foreign investment. Second, “responsibility” may be thought of in several different ways: (Level 1) responsibility for policy, standards and oversight, (Level 2) responsibility for provision (financing) and administration, and (Level 3) responsibility for production and administration. To some extent, shared responsibility among levels of government may be divided along these lines. With respect to primary and secondary education, for example, it is not uncommon for a national Ministry of Education to be responsible for determining national standards to be met by schools, teachers, and students, while regional (state) governments have the primary responsibility for financing education, and local (municipal) governments
Roy Bahl and Richard Bird

(or special school authorities) are in charge of delivering education. In Chile, for instance, although municipal governments deliver both basic health services and primary and secondary education, these services are largely financed by the national government, which is also responsible for ensuring that minimum national standards are met in all municipalities.

Applying this line of thinking to the project cycle commonly associated with infrastructure investment, one might end up placing the policy responsibility for ensuring that appropriate technical standards are met with one level of government, the primary regulatory (and likely some financing) responsibility with another, and the responsibility for actually delivering such services with a third. In the case of water, sewerage, and refuse collection, for example, one level of government might have responsibility for project design and management of water quality and waste disposal projects, another level of government with making provision for adequate water quality and quantity as well as environmentally sound wastewater and disposal services, and a third with operating the delivery system, including waste collection and disposal, and maintenance of the water-sewer infrastructure. In their attempt to develop a general framework, Boadway and Shah (2009) argue for the assignment of water, sewer, refuse, and fire protection entirely to local governments (the lowest level).

![Table 2. The Traditional Theory of Infrastructure Expenditure Assignment: An Illustration](image)

<table>
<thead>
<tr>
<th>Function</th>
<th>Policy, standards, oversight</th>
<th>Provision, administration</th>
<th>Production, distribution</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water, sewer, refuse, fire protection</td>
<td>L</td>
<td>L</td>
<td>L,P</td>
<td>Primarily local benefits</td>
</tr>
<tr>
<td>Police</td>
<td>S,L</td>
<td>S,L</td>
<td>S,L</td>
<td>Primarily local benefits</td>
</tr>
<tr>
<td>Education, health, social welfare</td>
<td>N,S,L</td>
<td>S,L</td>
<td>S,L,P</td>
<td>Transfers in kind</td>
</tr>
<tr>
<td>Natural resources</td>
<td>N</td>
<td>N,S,L</td>
<td>N,S,L,P</td>
<td>Promotes regional equity, common market</td>
</tr>
<tr>
<td>Fiscal policy</td>
<td>N</td>
<td>N,S,L</td>
<td>N,S,L,P</td>
<td>Coordination possible</td>
</tr>
<tr>
<td>Regulation</td>
<td>N</td>
<td>N,S,L</td>
<td>N,S,L,P</td>
<td>Internal common market</td>
</tr>
<tr>
<td>Foreign direct investment</td>
<td>N,L</td>
<td>L</td>
<td>P</td>
<td>Local infrastructure critical</td>
</tr>
</tbody>
</table>

Source and Notes: Extracted (including comments) from Boadway and Shah 2009, 134-35. N=national government, S=state (provincial, regional) government, L=local government, P=private or non-governmental.
but suggest that responsibility for policy and service delivery for other expenditure functions should be divided among governments as illustrated in Table 2.6

In practice, the unbundling of expenditure responsibility differs from country to country. Expenditure unbundling is common, but not a foolproof solution to the assignment problem. Most countries provide a central and a local government list of service responsibilities, sometimes in great detail. However, there often is a concurrent list that makes matters murky. The result can be overlap and confusion about who is responsible for what. For example, rural local government welfare programs in Karnataka state in India are overlapped by eight central government programs and 47 state government programs (Rao, Nath, and Vani 2004).

Three aspects of this brief summary of the traditional approach deserve close attention. First, the argument for local control depends not only on the beneficiaries being local, as emphasized in Table 2, but also on those exercising local control being more likely to meet local needs and, ideally, on local users’ ability and willingness to pay the costs. Only when local governance institutions ensure in an accountable way that those who benefit and pay are also those who decide what is done should all aspects of infrastructure decisions be assigned to local governments. Strictly earmarked user-charge financing (including an appropriate intertemporal dimension) meets this standard, but general fund financing, even if entirely local, will do so only when governance institutions are both strong and fully accountable. Many developing countries do not have such institutions—or perhaps, in some cases, may have conflicting local institutions (e.g., traditional, appointed, and elected officials)—so even when all benefits of public-sector activity are realized (and even paid for) locally, it is not always as clear as the conventional approach suggests that all aspects of the provision of such services are necessarily best provided locally.

Second, with respect to the important “shared” infrastructure responsibilities included in Table 2, the appropriate local component depends on clearly understanding who benefits, by how much, and when—and again, it is critical to tie together the financing and expenditure sides of the infrastructure budget (including operations and maintenance). Not only does this require careful consideration of spillovers or externalities (network, interurban, etc.), but there must be adequate horizontal and vertical coordination, cooperation, and burden-sharing among affected jurisdictions. Similar concerns often arise even with critical aspects of “local” activities such as water, sewers, and refuse disposal as well as “point” investments such as new schools or health centres. Few developing countries are likely to satisfy these conditions, so the guidance provided by theory may not always be as clear as the traditional literature suggests.

6. They recognize, of course, that in some cases it may be more efficient to contract out many aspects of the construction and management of projects as well as service delivery itself to private providers, although such contracting may increase the regulatory—and perhaps, in some instances, even the financial—burden on the relevant governments.
Infrastructure projects can never be lumped in one group and treated symmetrically and in the same way. Even with respect to economic infrastructure, for instance, big projects such as intercity highways are quite different from small ones such as village-to-market roads, not only in scale and complexity of design (as well as the need to fit the project within a national perspective), but also in financing, execution, and operation (maintenance). Small “purely local” projects can and should be left primarily to local decision-makers to decide on and to fund (including funding from general transfers and borrowing). The latitude for error (e.g., corruption) may be greater with such small projects, but the consequences of such errors are (and should be) largely borne by those who make them. Much the same may be true with respect to “point” projects such as the design and location of a particular health centre or school. Mistakes can and will be made; so long as those who make them bear the costs, they have a strong incentive to get it right. Large network projects, on the other hand, both because of their broader externality effects and because of the sheer magnitude of their technical and financial demands, usually require both support and guidance either from higher levels of government or from cooperating jurisdictions at the same level (although the appropriate design of such intergovernmental agreements is beyond the scope of this paper).

Third, even with respect to purely local infrastructure, localities require an ample tax base or ample access to capital markets or both to assemble the lump sums needed for most infrastructure projects. By definition, infrastructure has an important temporal dimension and capital financing must be timed and provided in the right way if the right projects are to be carried out in the right places. This is true even when in principle it should be solely local governments that decide what is “right,” albeit perhaps subject to standards of quality, quantity, and access established by higher levels of government and presumably also to monitoring and evaluation by them. Once again, few developing countries provide a promising environment for the effective implementation of such ideas.

The solution of the assignment problem with respect to infrastructure decentralization, as suggested in the traditional literature, is thus simple in principle but considerably more complex in practice. The simple principle is what is sometimes called “subsidiarity.” Decentralized provision of the infrastructure needed to fulfil local expenditure responsibilities should clearly be the responsibility of local governments. In some instances, restructuring subnational government borders or reassigning functions to different levels of subnational government may help achieve efficient decentralization. In others, the most efficient way to proceed may be establishing some sort of “special district,” although, as noted below, the effective governance and operation of such special-purpose governmental units can be surprisingly difficult to resolve (Berry 2009).

An additional complication arises because, in principle, since capital assets usually provide services over long periods, the most efficient and equitable way to finance such infrastructure is often through borrowing. However, in practice in most developing countries, local borrowing is subject to various regulatory and
quantitative restraints, especially when combined with the less developed capital markets in most countries. Much local infrastructure thus ends up being financed by capital grants (or subsidized loans) from higher levels of government, or even by foreign donors.

More broadly, one cannot consider decentralization and infrastructure in isolation from the critical issues raised in the “second-generation” fiscal federalism literature with respect to decentralized governance in general (Oates 2005; Weingast 2009). As this literature shows, local responsibility and accountability is crucial in any decentralization program if it is to be compatible with incentives. Without good local governance in a good intergovernmental system, local provision of infrastructure is likely to be far from the theoretical ideal. The appropriate long-term solution is not, however, to shift responsibility to levels of government that are unlikely to deliver anything better—and may even, from the perspective of local consumers, do worse—but to improve local governance institutions by establishing workable and meaningful local accountability by fiscally responsible governments. To do this, local governments need to have control over adequate resources for which they are fully accountable to those who provide them, whether local citizens through user charges and local taxes (including those servicing capital debts), or higher levels of government that provide grants or subsidized loans, or private financial institutions that provide infrastructure loans. Even if decentralization is perfectly designed, some failures—breakdowns of service delivery, payment arrears, even extreme insolvency—may occur, so that there is always a need not only for the central government to maintain a sound macroeconomic framework (e.g., with respect to subnational borrowing) but also to develop and, if necessary, implement, the equivalent of a “bankruptcy” takeover of failed local governments.

The requisite balance between assigning expenditure responsibilities to a decentralized governance institution that is just large enough to encompass the immediately relevant externalities while at the same time taking adequately into account the interests of all those affected (e.g., those in the areas where the refuse is dumped, buried, or burned) rather than just the interests of such specific groups as shareholders or public-sector workers is not easy to attain. Finding an expenditure delivery arrangement that passes this test becomes even more difficult when it becomes necessary to align “who pays” with “who benefits.”

Two other factors drive a wedge between the theory of assigning infrastructure responsibility and the practice. One is the inherent complication of the matter. Imagine trying to unbundle education expenditures into various activities, and then developing an operating definition of an activity such as “standard setting.” And then there are requirements that the benefit zone for certain infrastructure services be defined. Finally, there is politics. Unbundling may be a good idea for finding the right level of government to deliver a service, but elected politicians may be loath to give up control and may impose new mandates in the name of more efficient infrastructure service delivery.
2.2 From principle to practice

The traditional assignment of expenditure responsibility suggests that much of the spending on infrastructure will be made by local governments. Worldwide, about one billion people lack clean water and perhaps three billion lack access to adequate sanitation facilities. In principle, the annual infrastructure expenditure required to meet this one local need has been estimated at 2.0 percent of GDP in Sub-Saharan Africa and 1.7 percent in South Asia, with about 40 percent of these amounts required for new investment and the balance for operation and maintenance (Estache 2010). In practice, however, as Table 1 suggests for some Latin American countries, total local expenditures on all infrastructure have sometimes been less than these levels, and the share financed from “own” revenues is smaller.

It is not surprising that in practice, new investments made in even the most local of infrastructure in poor countries are largely financed by foreign aid and capital transfers. Indeed, in many cases local “ability to pay” is so low that even recurrent operating and maintenance expenditures, when made at all, are sometimes financed from such outside sources. 7

A major obstacle in the way of assigning more responsibility for infrastructure to local governments is that most local governments in developing countries lack the financial, managerial, and technical ability to do the job. Two solutions seem possible: kick it all upstairs to a higher-level government, or work away at the long, grinding process of building up local governance structures and managerial capacity. It is always quicker and easier to outsource tasks (e.g., through public-private partnerships) than to build up the institutional structure to enable local people to do the job themselves. Given the short time horizons under which aid agencies and national governments operate, they usually choose the easy way. In Bangladesh, for example, the national Local Engineering Department essentially centralized local investment. Although centralization probably improved the technical efficiency of projects and was usually enthusiastically supported by local governments glad to be freed of the need to tax their own people, following this path meant that local autonomy was significantly undermined. The result is that (1) local people have little or no ownership in the asset, so they tend to run it down rather than maintain it properly, (2) they still lack most of the capacities needed to do it right anyway, (3) the asset provided may not be what local people really want, and (4) the whole cycle is likely to continue as long as outsiders are willing to pay.

A similar system of centrally controlled infrastructure investment existed in Indonesia during the 1970–2000 period. However, this was replaced with a decentralized system in 2000, and the responsibility for choosing and implementing infrastructure projects began to shift to elected local governments.

Preferences

Preferences for capital projects will vary across regions within a country because of differences in the economic base, differences in the existing quality of
infrastructure services, and differences related to urban population densities. The idea that local familiarity and information can lead to better project selection and implementation is both generally accepted and supported by some country studies (e.g., Fiszbein 1997; Faguet 2004).

The demand for local control of infrastructure spending is high because the benefits are more readily identified by voters than is the case for many other government outlays: neighbourhood residents are more likely to care about better water supply or a new street lighting system than they are about an increase in public employee wages, the preparation of a new master plan, or improved tax administration. This kind of anecdotal reasoning leads some to the conclusion that decentralized infrastructure responsibility should improve allocative (economic) efficiency by matching provision more closely to local preferences. One example is some recent research showing that giving responsibility for water supply to local governments rather than public enterprises increased the quality of the output in Colombia (Bird 2012).

Other evidence suggests that local control may also lead to increased overall levels of infrastructure spending (Estache and Sinha 1995; Boadway and Shah 2009). A recent econometric study in Europe, for example, found that decentralizing revenue authority to subnational governments increased their investment in economic infrastructure, without reducing their “social” investment (Kappeler et al. 2012).

Although the political attraction of investing in highly visible public works is obvious and the economic outcomes of such projects may also be positive, legitimate questions may still be raised about the relative effectiveness of such programs in terms of poverty alleviation compared with, say, conditional cash transfers (McCord 2012). Even in this respect, however, some evidence suggests that appropriately locating such “geographic capital” (Jalan and Ravallion 2002) may have significant impacts on improving the access of the poor to markets, work opportunities, and welfare (Majumder 2012). As Fiszbein (1997) noted some years ago, for example, poor people in an isolated mountain village may quite rationally choose to invest resources in improving market roads rather than primary education—or whatever else planners in the capital city may think is best for them.

But are such potential efficiency gains from decentralization always—or even often—captured? There are a number of reasons why they may not be, starting with the possibility that the local population may not have a vote that allows them to express preferences (nor a credible exit threat), as in China, Vietnam, or Nepal. In China, for example, subnational governments have considerable discretion in making infrastructure decisions, but their political leaders are appointed and are accountable upward, not downward. Infrastructure accomplishments have been remarkable in China, but it is not clear how much of this is driven by local preferences (Dollar and Hofman 2008). Bird et al. (2011) discuss a study of about 500 villages that indicates that relatively little attention was paid to local preferences in making decisions about local expenditures in China.
Even where there is local voting, decision-making about public spending may be captured by powerful interest groups within the community, such as local politicians or community organizations, or by central ministries. The process of selecting and designing capital projects may not be weighted in favour of the preferences of the local population (Peterson and Muzzini 2005), or the procurement process may be corrupt and inefficient. Government structure itself may be an impediment where the large size of provinces and cities make the flow of benefits from capital investments all but invisible to individual households.

Some of these obstacles to capturing the potential efficiency gains from fiscal decentralization can be overcome or at least ameliorated. One way is to provide for submunicipalities within large urban areas and to give them some degree of decision-making power and a budget. This approach allows for implementation of small-scale projects at the neighbourhood level, and might provide a stronger voice for neighbourhood populations in the selection and design of larger capital projects. Examples of this are the barangay governments in metropolitan Manila and the extensive use of locally based benefit funding of investment projects in some Colombian cities (Bird 2012).

Another way to do a better job of capturing the benefits of decentralized infrastructure spending is to monitor client satisfaction with infrastructure services. The use of surveys to gain some information on citizen perceptions about the quality of services delivered can be an important input to identifying new infrastructure needs and maintenance requirements. While surveying is often done in the case of recurrent expenditures, it is less often done with respect to infrastructure (Peterson and Muzzini 2005).

Technical efficiency

Welfare gains from the decentralization of responsibility for infrastructure services can also result when services are produced at lower cost. One important way to lower costs is to deliver services more efficiently. A recent study in Ontario, Canada, for example, reports that 20 percent of treated water is “lost” in the distribution system before it reaches consumers (Herstein 2012). Such leakages through technical failures (as well as through theft) are often much larger in developing countries and might be lowered when providers are more directly

8. Berry (2009) provides a detailed account of the extent to which special-purpose local governments in the United States have been “captured” by such groups as developers, other business interests, and public-sector workers. The history of municipal infrastructure development around the world is replete with similar stories of elite capture: see, for example, the account in Briggs (1996) of city development in the United Kingdom in the 19th century, when most infrastructure was financed by (subsidized) private firms.

9. The barangay (district, ward, village) is the lowest level of local government in the Philippines. It is governed by an elected local council, responsible for certain local services and financed largely by a national transfer based on population and land area. There are over 1,700 barangays in the Metro Manila area.
responsible and responsive to consumers, although empirical evidence on this point is lacking.

One must also consider the technology used to produce and, especially, maintain, services. For instance, in rural areas, pumps that can be repaired locally may be preferable to more modern equipment that requires more expert servicing, just as schools that can be built with local labour and materials may be cheaper to construct and maintain than those that meet the latest specifications sent from the experts in the capital. On the other hand, sometimes infrastructure inputs may be more efficiently “outsourced” to central (or private) design, production and procurement, provided adequate attention is paid to local conditions and preferences.

An interesting recent project to improve public transit in Ahmedabad, India, by building a new Ahmedabad Bus Rapid Transit Service (ABRTS) illustrates some of these points. This project seems to be surprisingly successful in its first years (National Institute of Urban Affairs 2011) for a number of reasons. First, its designers paid close attention to experience elsewhere with similar projects, both successful (e.g., Bogotá) and unsuccessful (e.g., Delhi). Second, in both the design and the implementation phases, extensive efforts were made to consult not only with customers (actual and prospective), but also with such “stakeholders” as those who provide competitive modes of transport. As a result, substantial efforts were made to turn competing into complementary (feeder) modes. Third, close attention was paid to the specific urban context in which the system was to operate, by establishing routes that would reduce congestion, and minimizing conflict around religious sites. Fourth, an extensive “pilot” operation was carried out to test the system, and this led to significant improvements. For example, the trial buses had metal seats that (unsurprisingly) turned out to be too hot for comfort and were therefore replaced by others before the system was finally launched. Local people were involved in all stages of the project cycle, not just by securing their agreement and approval, but by actively involving them in identifying and initiating projects (what most needs to be done and where), in project design (exactly where and in what way), and even, to some extent, in execution and operation (using local labour where feasible), as well as in monitoring and evaluating services.

For capital-intensive services such as railways and national trunk roads, the presence of economies of scale rules out some functions for decentralized assignment. The cost of delivery by small-area governments would be prohibitive for such services. But for most capital projects, the right level to assign expenditure responsibility and the extent to which local governments are “right-sized” with respect to taking advantage of economies of scale (and scope) remains an open question (Fox and Gurley 2006).

Also debatable is the extent to which local governments take advantage of economies of scale.10 The empirical evidence is, at best, mixed. Byrnes and Dollery (2002), for example, reviewed research on economies of scale in the United

10. Much of the following discussion is based on Slack and Bird (2012).
Kingdom and the United States and concluded that only 8 percent of the studies found evidence of economies of scale in local government, 29 percent found evidence of U-shaped cost curves, 39 percent found no statistical relationship between per capita expenditure and population size, and 24 percent found evidence of diseconomies of scale. Studies that analyzed specific services (e.g., fire, housing) also showed mixed results. On the whole, there is no strong evidence of economies of scale with respect to most services for municipalities larger than 20,000 to 40,000 in population.\(^\text{11}\) This does not mean that there are no size effects or that per-capita expenditures do not fall as the scale of local government operations increases. More likely it reflects the difficulty of separating the pure effects of scale on government costs from the effects of everything else.

Economies of scale depend on the service in question and the units of measurement—such as the jurisdiction size or the size of the facility. Hirsch (1959), for example, estimated cost functions for police services, fire services, refuse collection, water, sewage, and education and found that expenditures per capita declined with the quantity provided for water and sewage (perhaps unsurprisingly, given the heavy infrastructure component of such services)—but that there was no similar decline for other urban services generally provided by local governments in the United States. For some services, expenditures per capita actually rose as output expanded, indicating diseconomies of scale. Other studies that have estimated cost functions have similarly found economies of scale for hard services such as water, sewers, and transportation, but generally not for soft services such as police, refuse collection, recreation, or planning (Bird and Slack 1993). Hard services are capital-intensive, so large government units can more readily make the substantial capital investments needed to extend the water distribution system or build a least-unit-cost-sized sewage treatment plant (Bahl and Linn 1992). Other services, such as policing, are highly labour-intensive and hence unlikely to show significant economies of scale. Presumably much the same can be said with respect to other labour-intensive services like social services, education, and, to some extent, even health.

\(^{11}\)There are problems with the methodology used to measure cost and output (Byrnes and Dollery 2002). In most studies, expenditures are used as the measure of cost and population is used as the proxy for size or scale. First, population may not be the best measure for this purpose. A larger population may mean greater need for expenditures, but the characteristics of the population will also influence need. For example, a municipality with a large proportion of elderly residents will have different expenditure needs from one with a younger population of the same size; an urban population will have different expenditure needs from a rural population. The density and geographic distribution of population may also be an important factor affecting both needs and costs. Second, population does not reflect the non-resident population that visits a local government area and uses its services. Third, with respect to the measurement of cost, expenditures are not always the best proxy because they not only include costs, but also reflect quality of services and possibly wasteful expenditures. Few studies of economies of scale include service levels.
Recent studies in Canada and Finland, for example, have found little evidence of economies of scale in large municipalities. Found (2012) analyzed economies of scale for fire and police in 445 municipalities in Ontario, Canada, from 2005 to 2008. He found that fire services exhibited U-shaped cost curves with a cost-minimizing population of approximately 20,000 residents. Police services also exhibited U-shaped cost curves with a cost-minimizing population of about 45,000 residents. In Finland, Moisio, Loikkanen, and Oulasvirta (2010) reported on a number of studies of the effects of municipal mergers on per-capita expenditures and found the results to be mixed, with the biggest cities showing relatively low cost efficiency with respect to basic welfare services. Other studies in Finland that focused on specific municipal services (health centres and schooling) found the optimal size of the municipality to be somewhere between 20,000 and 40,000 people (Moisio, Loikkanen, and Oulasvirta 2010).

As important as economies of scale (and scope) are economies of density (Bel forthcoming). For example, a recent study of annexation that analyzed 952 U.S. cities (with at least 10,000 people each) that annexed other municipalities between 1992 and 2002 found efficiencies from increasing land area, but only if the annexation was accompanied by higher densities (Edwards and Xiao 2009). If densities fall following annexation, per-capita expenditures may increase or decrease depending on the relationship of the change in land area to changes in density. The authors found that service delivery and administrative efficiencies are achieved with high-density developments, but compromised with low-density developments that are spread out and more costly to serve.

The question of economies of scale aside, there is some support for the argument that decentralization can lead to more cost-effective outcomes for capital projects. It can be argued that the cost of production is cheaper when local labour and materials are used and the bureaucratic costs of managing a project from the centre are avoided. There is some evidence to support this argument, though the results are far from conclusive (Peterson and Muzzini 2005). But this cost advantage does not hold in all countries. For example, the World Bank (2009a) reports that subnational government spending for roads in Colombia is expensive and wasteful. However, as Bird (2012) argues, it appears that this result reflects in part the unclear nature of the initial assignment of functions (which led to continual disputes about which level of government was responsible for what when it came to roads), and in part the confusing way in which roads were financed. In Colombia, all three levels of government are involved in financing this activity, but the two subnational levels receive most of their funding from the national government through several different channels. As a result, not only is the “ownership” of (and responsibility for) roads further confused, but the diverse and volatile set of funding sources makes it exceptionally difficult either to invest sensibly or to maintain investments properly once made.

Externalities
When the delivery of a service leads to impacts on households that reside outside the boundaries of the jurisdiction, the lower-tier governments will underspend (or
overspend) because they account only for local benefits and costs in their budgetary decisions. Because the social costs and benefits due to spillover effects are ignored, society will not achieve as high a level of welfare as would be the case if the service had been assigned to either the regional or even the national level. Functions with high spillover effects, such as public health services, are not good candidates for delivery by a subnational government unless the area served can be expanded so that the externalities can be internalized or the government can arrange for payment by non-residents that would compensate for the benefits received.

For purposes of working out the assignment of each function to different levels of government, it is necessary to estimate the benefit (cost) zone for each expenditure function. This is mission impossible because nearly all functions delivered by government impose external effects, yet rational public policy reform requires identification of the degree of the spillover effects and estimation of the costs of internalizing them. The analysis for each function is not likely to result in precise estimates for these benefit zones, but it may be good enough to guide the expenditure assignment decision for each infrastructure service.

The result of this exercise will be to assign responsibility for every infrastructure function to either local, regional, or central government.12 Sometimes, the external benefit or cost zone is so great that only central government responsibility will do. This is the case if the service benefits or burdens the entire national population (e.g., defence) or there is a strong interest in maintaining national standards (education or the social safety net). In some cases, the external effects are deemed small enough to be ignored.

As a rule, however, benefit and cost spillovers related to the provision of infrastructure services generally relate to an area smaller than the entire nation, especially in larger countries. Provincial- or state-level governments might be the right choice for delivery of functions such as intermunicipal roads and watershed management. Some externalities may be internalized by assigning functions to an even smaller jurisdiction: for example, water supply, mass transit, and urban planning might be best assigned to metropolitan-area-wide governments, although all too often such governments do not exist, have the wrong powers, or do not have appropriate boundaries (Bahl and Linn 1992).

Certain institutional arrangements may be used to address the spillover effects that come with decentralization.

• First, the process begins with a clear identification of services for which the externalities are too large to be assigned to subnational governments. Carrying out this task correctly is not easy because of the

12. Actually, if such an exercise were done seriously, it would likely end up prescribing a different geographic “benefit” level for every public-sector activity and subactivity—or perhaps even for every geographic locality. As noted earlier, however, such a proliferation of special service districts is seldom a good idea.
difficulty of separating true external costs and benefits from the desire of central ministries to continue their control over certain programs.

• Second, infrastructure responsibilities can be unbundled into subcomponents that do not have large spillovers and those that do. In China, for example, the central government sets standards for such services as health care, while the subnational governments are given responsibility for design and implementation of particular projects, and for financing.

• Third, central governments can use intergovernmental transfers of various kinds to induce subnational governments to correct for underspending on services with large externalities. As discussed further below, these usually take the form of conditional grants.

• Finally, it might be decided that the lowest levels of government cannot handle the externality (or even the management) issues. In this case, decentralization for certain functions is still possible in the form of special-purpose public enterprises that operate on a local basis. In some Indian metropolitan areas, certain local infrastructure services are delivered by single-purpose metropolitan special districts, which would seem to address the externalities and economies of scale issues. However, these parastatals are state-owned, and so the responsiveness to local preferences is weakened (Mohanty et al. 2008).

Another option is contracting for services, either to a private firm or to another local government. While arranging some form of contracting between larger and smaller local governments for service delivery is an attractive alternative in principle, in practice it has turned out to be difficult to sustain such arrangements, even in such well-governed and homogenous developed countries as Denmark and Finland.13

Managing infrastructure
A proper schedule of maintenance can markedly enhance the flow of services from capital assets and extend their useful life. But maintenance is problematic under a decentralized system, for three reasons. First, the base of local revenues that can be used to maintain the capital stock is limited. Subnational governments raise only about 2.5 percent of GDP from own sources in developing countries (see Table 3). Ingram, Liu, and Brandt (2013) estimate that the required annual maintenance costs for urban infrastructure in developing countries is equivalent to about 2 percent of GDP.14

13. See the extensive discussion of such problems in Kim, Lotz, and Mau (2010).
14. An earlier study (covering rural and urban areas in developing countries) for the 2005-2010 period estimated future maintenance needs in developing countries to average about 3.3 to 3.5 percent of GDP (Estache 2006).
Second, local officials may be prone to ignore maintenance in favour of more high-profile construction of new infrastructure or in favour of satisfying such other demands as public employee wage increases. Moreover, from their perspective it may be politically simpler to obtain new capital financing from higher levels of government than to obtain more tax revenue from their constituents.

Third, as mentioned earlier with respect to roads in Colombia, the ownership of the public asset should be clear, lest there be confusion over which level of government is responsible for its maintenance (World Bank 2009a).

Are there institutional and behavioural arrangements that can overcome the production efficiency disadvantages of subnational governments in the decentralization of infrastructure services? Four possibilities might be suggested.

• First, if a service has been assigned to subnational governments, limit interference from upper-level governments that may increase costs. The involvement of higher-level governments (for example, by imposing mandates) should be limited to concerns about externalities and perhaps the regular monitoring of maintenance. Mandates to spend a specified minimum share of the budget on economic development, or a specified maximum share on personnel, reduce local spending flexibility if they are effective and are not easily enforced in any case. Some countries (such as Colombia) commonly earmark significant portions of general intergovernmental transfers to “investment,” largely to avoid the dissipation of such transfers in current expenditures—which all too often line the pockets of local officials and their friends and relatives rather than provide valued services to citizens in general. It is far from clear, however, that such earmarking achieves its objective. Indeed, as noted later, there are both theoretical and empirical support for the contrary view that reducing earmarking is more likely to improve than to distort the allocation of local resources.

<table>
<thead>
<tr>
<th>Region</th>
<th>Subnational government expenditures</th>
<th>Subnational government taxes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Percent of total government expenditures</td>
<td>Percent of GDP</td>
</tr>
<tr>
<td>Developing countries</td>
<td>18.8 (16)</td>
<td>5.1 (20)</td>
</tr>
<tr>
<td>Industrial countries</td>
<td>27.8 (26)</td>
<td>13.9 (26)</td>
</tr>
</tbody>
</table>

Source: IMF (various years) and estimates drawn from case studies.
Note: Data reported are unweighted averages for the 2000s for years in which data are reported.
*The number in parenthesis shows the number of countries included in the comparison.
Second, in the case of urban governments, expand the tax and user-benefit charge powers of subnational governments. This involves giving subnational governments access to a larger tax base and the autonomy to use this base. Increased provincial and local revenues would provide a funding base for infrastructure maintenance.

Third, close the “back door” to infrastructure finance that is often provided by higher-level governments. In Mexico, for example, as much as 10 percent of intergovernmental transfers to subnational governments are given in the form of ad hoc grants to subnational governments (Bahl and Sethi 2012). This approach discourages the mobilization of local revenue and reduces the transparency of the system of intergovernmental transfers.

Fourth, unbundle infrastructure services so that subfunctions not characterized by clear advantages of centralization can be assigned to subnational governments. Costs can be reduced by disaggregating infrastructure expenditures into components, and making assignments on a basis of comparative advantage. For example, while technical specifications required to ensure water quality may be a matter of national concern, the construction of major water supply and sewage lines (like interurban highways) may best be handled at the regional level, and local distribution lines could be the responsibility of local governments. Such decentralization is likely to work better when central governments make an effort to establish good “framework” laws (e.g., on tendering and on intergovernmental agreements), upgrade local capacity to deal with such issues, and monitor and evaluate outcomes. In a sense, the issue here is similar to that raised below with respect to public-private infrastructure projects: in order to reap the full benefits of decentralization, investment must be made in improving the regulatory framework at the national level and executive capacity at the local level.

Corruption

A possible cost of fiscal decentralization is that it may lead to greater corruption. The thinking here is that the “closeness” between elected local politicians and the local political power structure breeds corruption. Both Prud’homme (1995) and Tanzi (1996) have presented this view as one of the dangers of decentralization. There are further reasons why one might expect more corruption in a decentralized fiscal system. One is that the probability of successful stealing is increased by the weakening of central authority and monitoring. Various students of corruption have placed the blame on the greater number of contacts with public officials in developing countries, on lower-paid local public officials who have more incentive to steal than higher-paid central officials, and on local government voters who have not yet learned to use their power to monitor and discipline their employees. These problems may be especially serious with respect to infrastructure, where there is
more latitude for fraud, bribery, embezzlement, and patronage than with respect to other, more “regular” activities of local public officials and politicians. Moreover, local corruption, even if smaller in scale, may be particularly damaging to building “trust” in government, owing to the more visible inequalities that may result.

Nonetheless, the case that corruption and decentralization are positively linked might be more impressionistic than real, and the empirical evidence is divided on whether corruption costs are greater under a more decentralized system.\(^{15}\) Perhaps the common perception of high local corruption is unduly influenced by its greater visibility; corruption may be even greater with respect to centralized decisions if they are less transparent.

We are not aware of comparable recent estimates of the cost of corruption in the infrastructure sector. Data are scarce and the conceptual model is not easily worked out. In earlier work, however, Tanzi and Davoodi (2000) argued that corruption will lower infrastructure spending (fewer projects will be undertaken), while Mauro (1995) argued that corruption is more likely to raise infrastructure spending (higher unit costs). Whatever the outcome, information is scarce and unreliable, and helpful and feasible solutions short of a fundamental revision of the relationship between state and citizens are not easy to design or implement. Politicians and officials who gain much of their income from exploiting their monopoly power to grant licences, bestow contracts, or provide services are not likely to give it up easily. Neither regulation nor privatization seems to provide complete answers to controlling the problem.

As Estache (2006) notes, corruption is a symptom of a deeper underlying problem—the lack of political commitment and accountability. For example, Bardhan and Mookherjee (2006) have shown that simply financing local infrastructure through user fees rather than local taxes or intergovernmental transfers will reduce corruption, no matter how poorly local democracy works. As we discuss below, however, almost no infrastructure investment in developing countries is financed this way, and this situation seems unlikely to change soon.

**Capacity**

Subnational governments in developing countries often do not have the ability to design, build, and operate infrastructure, and seldom know enough to “outsource” the needed skills efficiently and effectively. Many who write on the possible dangers of decentralization emphasize the poor quality of local administration in most developing countries (Prud'homme 1995), although the variability even within specific countries makes generalization difficult. On the other hand, some evidence suggests that the quality of municipal management of public service delivery has improved in recent years in certain developing countries (World Bank 2009).

While there is often good reason for concern about the capacity of subnational governments to deliver services, to a considerable extent countries get the local

\(^{15}\) For reviews of this literature, see Martinez-Vazquez, Arze del Granado, and Boex (2007), and Boadway and Shah (2009).
governments they want and deserve. Subnational politicians and officials, like those at the central government level, respond to the incentives they face. If those incentives discourage initiative and reward inefficiency and corruption, it should not be surprising when local governments turn out to be corrupt and inefficient. 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 sometimes demonstrated significant improvements in administrative capacity within a relatively short time (Fiszbein 1997; Faguet 2004).

As Fiszbein (1997) noted in an early evaluation of decentralization in Colombia, for example, when given the chance to do more when they received more resources after a constitutional revision, some—not all—municipalities took steps to improve their capacities. Some, for example, improved the skills of local officials through competitive hiring, some shared the services of professional staff with neighbouring municipalities, and some undertook more training of municipal employees. Some municipalities also improved their capacity to carry out effective infrastructure projects. One, for example, privatized road maintenance; another put private developers in charge of the construction of urban roads; others introduced computers to monitor water and sanitation services, shared equipment with others, and directly attempted to improve their ability to manage municipal projects. While by no means all did such things, one result of the increased local engagement and “ownership” was that surveys found most respondents trusted the local government more than the national government to deliver the goods and services that they wanted (Fiszbein 1997).

Most municipalities concentrated on roads, education, and water works. These priorities may not have been what the central officials previously in charge thought was most important, but these were the needs local people perceived, and they were the needs that at least some of the newly empowered and responsive local governments attempted to meet. Although over the succeeding decade of political turmoil and civil strife, much of this initial positive response appears to have faded away, even a decade later, surveys continued to show that most Colombians were happier to pay local than national taxes, presumably because they felt they were getting more for their money (Acosta and Bird 2005).

3. Reconciling Theory and Practice

Theory tells us that efficient infrastructure is often decentralized infrastructure and that decentralized infrastructure should be efficient. In practice, however, the underlying conditions needed to ensure that this proposition holds are seldom met. The result is that in too many instances not only is infrastructure not provided efficiently, but the perceived failures of decentralized infrastructure lead to “corrective” actions that prevent decentralization itself from realizing its potential advantages. In most cases, the answer to these problems is to “decentralize” correctly, which means, essentially, satisfying three conditions:
• First, public-sector functions must be sufficiently unbundled to assign them to the correct levels of government: both who is responsible for what and who is accountable to whom for what must be clearly delineated.\textsuperscript{16} Clarity in assignment must be matched by accountability, in terms of both 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.

• Second, subnational governments must be adequately accountable in political, administrative, and financial terms both to those they are supposed to serve—their residents—as well as to those above them in the governmental hierarchy who may be responsible for developing policy, regulating how it is carried out, and often financing the activities of such governments. Such “dual accountability” is neither easy to design nor to implement.

• Third, the achievement of both the previous conditions is premised on the design and implementation of a sound intergovernmental finance system, one that devolves appropriate revenue-raising powers to subnational governments, makes it possible for them to employ these powers effectively, supplements local finance when appropriate through well-designed transfers, and provides adequate access to private sources of finance, particularly with respect to financing infrastructure investment. We develop some aspects of this third point in the present section.

3.1 Improve local governance

The efficiency of decentralized expenditure assignment for infrastructure is enhanced by local financing of these capital projects. If full financing is by transfers from higher-level governments and loans from state banks, then local officials are not fully accountable to the voters for the quality of services delivered.

Subnational governments in developing countries have access to three sources for financing capital expenditures: own source taxes, charges, and other non-tax revenues; intergovernmental transfers; and various types of privatization schemes.\textsuperscript{17} In many countries, however, both the structure of these financing instruments and their implementation impede more efficient and equitable outcomes in the delivery of infrastructure services.

\textsuperscript{16} As mentioned earlier, experience suggests that it is important not to separate responsibilities for building infrastructure from those for operating and maintaining it in order to avoid creating distortionary incentives affecting both current and capital expenditures.

\textsuperscript{17} Countries may also receive capital transfers from external donors. As Estache (2010) notes, aid-financed infrastructure is especially important in some countries in Sub-Saharan Africa. However, we do not consider this source further in this paper; for a review, see Kharas and Linn (2013).
No unified theory of revenue assignment exists that will allow identification of both the right level of subnational government financing and the right mix of revenue instruments to use (Martínez-Vázquez, 2013). As Bardhan and Mookherjee (2006) show, if infrastructure investment is financed by user charges, it is efficient. More generally, to the extent that subnational government revenues are raised according to the benefit principle—so that at the margin those who benefit from local infrastructure investment are those who finance it—the financing of infrastructure will be efficient. If locally raised taxes are paid entirely by those who benefit from local services, subnational governments will not overspend because they will not be able to export part or all of their tax burden to other jurisdictions (McLure 1998).

This condition may seem strict, but there are plenty of viable revenue options for covering this amount, including property and non-property taxation, and user-benefit charges (Bahl and Bird 2008). Significant increases in the mobilization of subnational government revenue through such “good” (local benefit) taxes and charges are feasible in many countries. However, higher-level governments in developing countries are seldom willing to give up productive revenue sources such as income or general sales taxes to subnational governments. Already strapped for revenues, they are hesitant to introduce competition for their tax bases, especially in the large urban areas where most taxes are raised, even when those are the subnational governments that both need and could use most effectively additional revenue from such sources.

Elected officials in some subnational governments are also hesitant to push for new taxing powers, precisely because it would ultimately force them into a position of greater accountability to their constituents. They prefer to claim increasing shares of national transfers (or foreign aid) rather than ask their residents to pay higher taxes. All too often, when subnational governments do get the green light on taxation, they are saddled with a limited choice of instruments and, in some instances, implement poorly structured taxes on business. These taxes may have undesirable efficiency effects but are nonetheless used extensively because subnational governments are able to impose them without having to rely on decisions by higher-level governments that would otherwise permit them to raise property taxes (Bird 2006; Martínez-Vázquez 2012).

### 3.2 User Charges

For utilities in particular, user charges are the obvious and ideal source of finance.\(^{18}\) Beneficiaries pay, and if the charges are set at the level of full cost recovery, debt service and maintenance can be supported. User financing of infrastructure ensures that those who benefit pay for what they receive, which is one of the best ways of ensuring that a particular service is worth what it costs. However, there are two big problems with advocating user-charge financing of infrastructure in developing countries.

---

18. For further discussion of the appropriate role and design of user charges, see Bahl and Linn (1992); Bird (2001).
One problem commonly perceived is that many people in developing countries are simply too poor to pay much. Too often this concern has resulted in too-low prices that benefit mainly richer consumers. Many governments grant subsidies and hold down tariff rates, with the result that much less than full-cost recovery—often even less than operating-cost recovery—is the rule rather than the exception (Estache 2010). The bulk of the effective public subsidy is captured by the few and fortunate and the resulting financial stringency ensures that not only do the poor continue to go without access to essential services such as basic water and sewage facilities, but also that the services provided are inadequately maintained, with resultant losses in supply. Many schemes to overcome such problems have been put forward over the years, but few countries appear to have made a serious attempt to price public services efficiently while taking adequate account (e.g., through “lifeline” or other tariff structures) of the need to provide a basic level of service to those who are unable to bear the full cost. 19

The second problem with user charge financing is that no one likes it, especially when paying for what one gets is not now the norm. People who are now receiving (often poor) service do not see why they should have to pay more and are unlikely to believe that they will, someday, receive something better for their money. People who do not now receive services may not believe that if they agree to pay more than those now getting the service pay, they will get some benefits from doing so. Politicians do not want to make current service recipients unhappy by making them pay and cannot easily sell “pay for what you get” to people who have long seen others getting without paying. While some experience suggests neighbourhood groups and small villages can agree to finance small infrastructure works that will benefit them directly (Bird 1995), there is little or no evidence that anyone has been able to devise and implement a “corrective” user charge structure to rectify the inefficient, inequitable, and ineffective fee structures now found in most developing countries.

3.3 Property taxes
The property tax can be an efficient source of revenue for financing infrastructure. In many cases, it finances infrastructure that enhances property values, and its revenue potential is large. Bahl and Martinez-Vazquez (2008) estimate that the property tax yields 2.2 percent of GDP in revenues in industrial countries, but only about 0.66 percent in developing countries. When the family of taxes on real property is considered—transfer taxes and various forms of betterment levies—the revenue potential is even greater, although the economic effects of transfer taxes leave much to be desired. Moreover, the property tax has the virtue of being a levy that higher level governments do not want in their tax structure, either because they recognize the inherent advantage of local governments in administering the

19. For an early review, see Bird and Miller (1989). Unfortunately, little has changed for the better since that paper was written.
tax or because they recognize that the unpopularity of the tax goes well beyond its usually meagre revenue yield.

But property taxes do not finance much infrastructure. The political unpopularity of the residential tax often leads to a neutering of the tax base, sometimes by higher-level governments who take credit for new exemptions while leaving local governments to deal with covering the revenue costs. Moreover, administration is often weak and always expensive (Bird and Slack 2004; McCluskey and Franzsen 2013).

Subnational governments might also use property-based taxes to recapture the value from public investments, and in some cases this approach has yielded significant amounts of revenue (Alm 2010). In some countries, governments have been aggressive about using value-capture methods in connection with infrastructure finance (Colombia), whereas in others (Philippines), it remains a largely unused resource. But even in the best of cases, such as Colombia, making a land value recapture scheme successful requires a great deal of technical and political investment (Bird 2012).

### 3.4 Other taxes

Subnational governments in Argentina, Brazil, and Colombia have mobilized significant revenues with subnational government taxes. Argentinean and Colombian subnational governments use a form of business turnover tax, while Brazil taxes gross receipts from the sale of services. But these taxes are highly distortive and serve as general revenues rather than as earmarked sources of funds for infrastructure. Despite repeated calls to modernize these sales taxes or business taxes, they continue in their present form.

The other option that is often considered is to abolish distortive subnational government taxes. South Africa levied a combination gross-receipts-and-turnover tax, but this was abolished by Parliament in part because of its weak structure and administration. India's octroi, a duty levied on goods entering a city for sale, is still a primary source of revenues for the Mumbai municipal corporation (Pethe 2013) but has been abolished elsewhere in the subcontinent. In these cases, the abolished taxes have not been replaced with an equally elastic local revenue source.

Motor vehicle taxes seem a good candidate for infrastructure finance, particularly for roads and mass transit. All of the possible instruments—licences, motor fuels, parking fees, and tolls—can be structured as benefit taxes. Certainly the tax base is growing. It is pointed out that for India, the motor vehicle population increased one-hundred-fold between 1951 and 2004 and at an even faster rate in the 2000s (High Powered Expert Committee 2011, 56). On the administration side, the most difficult problem is to allocate the funds to the place where the vehicles are used rather than where the distributor is located, but this might be done using distributors' shipment records until collection at the pump becomes a possibility.
3.5 Non-tax revenues

Like motor vehicle taxes, some non-tax revenues might be earmarked for infrastructure finance, even though the link between beneficiaries and burdens might not be so clear. Chinese metropolitan governments have been particularly innovative and have engaged heavily in land sales (long-term leases) as a method of mobilizing resources for infrastructure finances. For all local governments in China, land leases now account for about 30 percent of revenues (Wong 2013). Land sales have great advantages, namely the revenue potential and the low political cost (at least in China) of raising money this way. But even in a unique setting like China, there are drawbacks. Land revenues are sensitive to the real estate cycle and land value collateral for loans is risky; “easy money” can lead to overspending in local government budgets; the opportunity costs of converting land to urban use can be underestimated; and government-owned land is an exhaustible resource.

A more difficult arrangement to assess is the earmarking of certain revenues for infrastructure development. An example is Peru, where natural resource revenues are distributed on a derivation basis and must be spent for economic development purposes (Canavire-Bacarreza et al. 2012; Martinez-Vazquez 2012). On the one hand, the goal of such a program is to replace the lost “heritage” of the region with a new infrastructure. Management issues aside, there is some merit to this justification. On the other, this kind of earmarking introduces a function-follows-finance model and provides little incentive for efficient spending.

Even without a separate system of capital transfers, some countries (e.g., Colombia) commonly earmark significant portions of general intergovernmental transfers to “investment,” apparently in large part to restrain the feared dissipation of such transfers in current expenditures. It is far from clear that such earmarking achieves its objective. Indeed, some evidence suggests that reducing earmarking is more likely to improve than to distort the allocation of local resources. The theoretical argument is simply the standard decentralization theorem: that allowing funds to be spent according to local tastes yields, on the whole, superior allocative and distributional results. Empirical evidence in Norway (Borge et al. 2012) supports this argument strongly.

Although one may doubt the relevance of this experience to developing countries in which “elite capture” of local governments seems likely, studies in Bolivia and Colombia (Faguet 2005) also point in this direction. Most interestingly, perhaps, a recent study of the European Union (Kappeler et al. 2012) finds that revenue decentralization tends to increase subnational infrastructure investment. Earmarking local revenues for infrastructure may make sense when there are good reasons (efficiency, equity, and management) for doing so. Small projects with well-defined benefit groups are most likely to meet these conditions. In other cases, 20.

20. In Colombia, the potentially pernicious effects of such earmarking are perhaps mitigated to some extent by the fact that “investment” is interpreted to include so-called “social investment” in health, education, and so on.
however, such earmarking may distort local preferences, exacerbate perverse incentives to build new rather than maintain existing infrastructure, and connect revenue sources with particular expenditures in ways that lack economic and political logic. And, finally, there is the fungibility problem: it is not usually possible to separate the expenditure of the earmarked revenues from the expenditure of other revenues.

3.6 Intergovernmental transfers
Since subnational governments in most developing countries have limited own resources and little access to private capital markets, to carry out costly public works they usually rely heavily on grants (or subsidized loans) from higher-level governments. Depending on how the transfer system is structured, it may have a direct or an indirect effect on infrastructure finance.

Most countries structure their transfer systems using some combination of three approaches. The first is unconditional grants, which are an indirect way to provide infrastructure finance. These funds can be used for maintenance of the public capital stock and for debt repayment, though the grant funds are not directly tied to either.

In the eyes of lenders, the security associated with unconditional grants as the base for repayment is not strong. Some countries attempt to get around this problem by ensuring that these transfers cover debt service costs. This means providing for an intercept arrangement or allowing these transfers to be pledged to repay debt, as has been done in Mexico (Revilla 2012). Although many countries prohibit such arrangements, this approach to, in effect, “earmarking” grants to service the debt on capital projects deserves consideration.

A second approach is to make grants to local governments conditional on their expenditure for a particular capital purpose, often with no matching arrangement and in practice often with little supervision, either before or after the money flows. Often these conditional grants are distributed on a formula basis, and in other cases they take the form of cost reimbursements.

Capital grants to support local infrastructure projects, sometimes with a requirement of matching funding from local sources, are not uncommon, but the appropriate design of such transfers is not always easy to determine. A number of questions need to be considered: Is one aim to rectify imbalances in the distribution of existing infrastructure (Ahmad and Searle 2006)? To what extent are such transfers intended to improve economic efficiency—as is presumably the case with respect to most economic infrastructure such as transportation—or to ensure equalization by, for example, ensuring at least a minimum standard of such public services as education and health by providing hospitals and schools in particular areas (Josie et al. 2008)? Whatever the intended goal, in allocating such transfers between jurisdictions and in determining the appropriate matching rate, the nature of capital spending requires that the concept of fiscal capacity be expanded to encompass not just tax capacity, but also the ability to access credit markets (Herrero-Alcalde et al. 2011).
The third approach is more directly tied to infrastructure budgets. It can take the form of a block grant for infrastructure services on a matching basis. In India, for example, a large federal grant for urban infrastructure development and slum upgrading is allocated to cities on a matching basis (High Powered Expert Committee 2011). The program was introduced in 2005, and while it has succeeded in focusing increased attention on urban infrastructure issues, implementation has been slow (Rao and Bird 2011). South Africa uses a more formal municipal infrastructure grant, designed primarily to improve services in poor neighbourhoods (van Rynneveld 2007). In Brazil, ad hoc grants are made to support specific projects.

If done in conjunction with increased local revenue mobilization and cost-recovery levels of user charges, a case can be made that any of these approaches to intergovernmental transfers can lead to a more efficient delivery of infrastructure services by local governments. Still, conditional grants can be improved in most countries in a number of ways. For example, the terms and conditions of such transfers might require subnational governments to prepare adequate investment and maintenance plans, as well as an appropriate user charge policy. The governments receiving such transfers should be selected by a systematic process that pays attention to both need and capacity factors and to the economic evaluation (cost-benefit analysis) of the project in question. Technical assistance should be made available to subnational governments to permit them to develop plans, arrange financing, manage construction, and operate the facility (or to contract out its operation) efficiently. Finally, to ensure accountability as well as good outcomes, the execution and operation of the grant-aided work should be monitored and evaluated, with periodic progress reports, field inspections, and formal evaluations of outcomes, including consumer surveys.

These conditions may be too much to expect in many developing countries. Nonetheless, as the early experience with “municipal development funds” (Davey 1988) has demonstrated, if such conditions are not satisfied, the results of capital transfers and loans are unlikely to live up to expectations. Subnational infrastructure finance inevitably depends greatly on transfers from higher-level governments. These transfers may be efficiency-enhancing when they take the form of conditional grants designed to correct for underspending on services characterized by externalities. Few such transfers are found in practice, however, partly because the design of a conditional grant for infrastructure involves guesswork if (as theory suggests) one goal is to correct for underspending as a result of not taking externalities into account. It involves identifying a target for the optimal level of spending for the function that takes account of social as well as local benefits, as well as having some notion of the elasticity of demand for the infrastructure service. Careful targeting along these lines is an essential ingredient of effective subsidy policy in resource-constrained countries, but most infrastructure-related transfers, perhaps understandably, take a more general approach.
3.7 Borrowing

Borrowing is not itself a source of revenue, but rather an efficient way to arrange payment for the purchase of public assets that have a long life. By matching payment for the infrastructure with the pattern of consumption of the asset, governments can immediately begin capturing the returns from infrastructure investments while deferring the payments. But in the end, the loans must be repaid—from the own-source revenues of the subnational governments, from user charges, or from intergovernmental transfers.

The essential foundation of a sound subnational borrowing program is thus a sound subnational fiscal structure in terms both of access to own-source revenues and a well-designed and stable intergovernmental transfer system. Still, loans are often essential because they offer subnational governments a way to attract the large amount of funds necessary for financing the construction phase. Moreover, in principle they are an economically efficient (and equitable) source of finance because the debt has to be repaid (from current revenues) during the life of the project, that is, while it is providing benefits to local residents.

But borrowing by subnational governments can lead to problems, as students of local public finance have long argued (Prud’homme 1995; Tanzi 1996). The revenue stream of local government revenues may not be large enough to sustain repayment, but borrowing occurs anyway in anticipation of some form of bailout. This well-known moral hazard problem has led to overborrowing and to some form of bailout of subnational governments in Brazil, Argentina, South Africa, and more recently in China (de Mello 2007; Wong 2013). Many countries attempt to control overborrowing by subnational governments with various forms of fiscal responsibility legislation (Liu and Webb 2011). These programs have met with varying degrees of success. The problem with bailouts is that subnational governments learn that there is always a way around the hard budget constraint, and they may return to the strategy of overspending on infrastructure. A good borrowing framework can head off this tendency, but it may close the borrowing window for those local governments that have the greatest infrastructure gaps.

Bundling subnational borrowing—for example, by having regional entities borrow for smaller local authorities—may make sense on cost grounds. Similarly, packaging subnational borrowing in some form such as a national infrastructure bank may make sense when financing flows to projects that, although the responsibility of subnational governments, are of national significance (e.g., as part of a national network of roads or electricity distribution). In most cases, however, loans from such public financial institutions are extended on subsidized terms and are in effect a variant form of matching grant.

Although the lessons of unfortunate experiences with such “soft” funding may have been learned, as yet there appears to be no foolproof institutional way to avoid the hard necessity of first establishing subnational finances on a sound and sustainable basis so that at least some of them—probably the largest and better-off—are, when financial markets are sufficiently developed, creditable candidates for private-sector financing. When subnational governments are charged with an
important role in providing infrastructure, many of them will need both technical and financial support to play that role adequately, but it is important to ensure that financial resources do not arrive in ways that distort the incentives facing subnational decision-makers in ways that may damage both the provision of infrastructure and, more broadly, governance at the subnational level.

### 3.8 Public-private partnerships (PPPs)

Much the same can be said with respect to PPPs. Over the past two decades, many decision-makers have promoted increased private involvement in infrastructure creation to increase the efficiency of service provision and to provide badly needed resources to support urban infrastructure investment. Klein (2012) argues that both of these justifications need to be qualified. Private-sector “resources” come from user charges, which can also be realized by local governments if levied at cost-recovery rates. And, the private-sector “expertise” may be a step up from local government capacity but comes with a higher cost of financing.

In addition, there are questions about how PPPs deliver services: full privatization with various degrees of regulation, or some form of contracting for operation? The build-operate version is attractive in part precisely because it offers a way to get facilities built without incurring highly visible government debt.

It would be fair to say that the high hopes for PPPs in the 1990s and early 2000s never materialized. As yet, PPPs have added relatively little to urban capital financing in developing countries (Alm 2010; Annez 2007). Moreover, less than 10 percent of the investment has been in the high-priority water/sewer sector, mostly because of pricing risks, and an even smaller share in the form of full or partial privatization (Menard forthcoming). Most PPPs have focused on the energy and telecommunications sectors (Klein 2012). Where private capital has been attracted to these areas (such as in Bolivia, Venezuela, and Argentina), the outcome has often been conflict and failure, leading to “re-nationalization” of projects initially carried out by PPPs.

Opinions differ sharply with respect to both the merits of PPP arrangements and how best to design and implement them. In a recent assessment of the literature, for example, Merk et al. (2012) conclude that a critical feature (if a PPP is to minimize project costs) is that the contract should be “global,” that is, a single contractor should be responsible for managing the whole project in order to reap full economies of scale and scope and provide maximum incentives to invest and innovate. In contrast, Siemiatycki and Friedman (2012), considering urban transit projects in which a major issue is how to allocate “ridership (demand) risk” argue that “unbundled” contracts that exclude facility operation are often preferable, to attract more competitive bids and to lower the cost of private-sector borrowing. The contradictory nature of these recent assessments highlights the importance of the specific context and institutional and regulatory settings (e.g., the relative

---

21. For a recent study emphasizing the importance of appropriate incentives, fiscal and otherwise, with respect to establishing and sustaining sound decentralized governance, see Faguet (2011).
development of financial markets and who sets user charges), as well as the detailed specifics of PPP contracts in determining outcomes.

Even when a transit project is carried out entirely by the public sector, similar regulatory and financial factors are critical. China, for example, has in recent years carried out huge urban infrastructure investments through what may perhaps be thought of as a Chinese variant of PPP, in which both parties to the contract are really “public.” To illustrate, the extensive Beijing metro system, like many other projects in China, was financed primarily by bank loans to a local investment corporation—the Beijing Infrastructure Investment Corporation—created and controlled by the local government (Su and Zhao 2006). Since the national government strongly encouraged local governments to make such investments and banks to finance them in doing so, presumably the banking sector considers such loans to be guaranteed by the state. In addition to this implicit subsidy, loans to local investment corporations have generally received an explicit subsidy in the form of an interest rate about 10 percent less than the normal rate of long-term debt. In reality, however, most local borrowing is directly serviced from local revenues, which depend on revenue from leasing land. Such revenue is highly sensitive both to property values and to the amount of land sold: in 2011, for instance, Beijing’s revenue from this source decreased by 36 percent from the 2010 level.

Even if one assumes that the whole rapid transit system has been and is optimally designed, constructed, and operated, there are still problems ahead. Current plans to expand the metro system further in the next few years and a flat-fare policy that already requires an annual operating subsidy (and will need an even larger one in an expanded system) threaten the sustainability of the financing model used to build not only the Beijing metro but also much of the extremely impressive development of urban infrastructure in China in recent years. This approach to building infrastructure is an example of what Wong (2013) calls “riding the tiger.” It may soon require reconsideration and in all likelihood substantial adjustment—unless the urban real-estate sector can realistically be expected to continue to boom at pre-2010 rates for the next few decades.

To take a different, and perhaps more widely applicable example, Ahmedabad’s Bus Rapid Transit System was 35 percent financed by a national urban development program (the Jawaharlal Nehru National Urban Renewal Mission, JNNURM), 15 percent by the State of Gujarat, and 50 percent by local sources (including a dedicated urban transport fund). The system was developed and implemented through nine separate PPP arrangements negotiated between the special public corporation (Ahmedabad Janmarg Limited) created for the purpose and various private providers. These arrangements covered everything from the major system investment (bus stations, bus corridors, and flyovers, as well as buses) to housekeeping and parking (National Institute of Urban Affairs 2011).

The resulting system has both substantially improved “people transport” in the city and has (like the somewhat similar earlier “Transmilenio” BRTS in Bogotá) won national and international acclaim as a model worthy of emulation elsewhere.
Some aspects—notably dedicated funding sources for urban transit systems—have already appeared in several North American cities in recent years: Vancouver, Chicago, Los Angeles, and Salt Lake City, for example; all collect fees and taxes dedicated to a regional transportation authority (Institute on Municipal Finance 2012).

There are, of course, risks associated with public-private partnerships. For the private sector, for example, there are risks that the regulatory framework or pricing commitments may change and cause delays in the project. Annez (2007) and Ingram, Liu, and Brandt (2013) argue that the inherent riskiness of urban investments in water and sanitation is the main constraint to increasing the flow of private capital. There is a weak record of full-cost recovery, and often an unwillingness of local governments to stand behind the kinds of tariff levels and regulatory arrangements necessary to attract private investors. In many countries there is what Pethe (2013) describes a “trust deficit” between public and private sectors that has resulted in episodes like the Bolivian “re-nationalization” mentioned above. For all these reasons, the World Bank (2009, 32) points to “few positive results” in efforts to attract private financing of municipal services.

For the public sector, there is the risk that the services provided may not be what the public wants. There is also the risk that the private partner will fail, or insist on renegotiating the contract, and the public sector will have to take on the obligation in full. How successful such arrangements are from the perspective of either partner depends very much on how the contractual arrangements are structured and how the risks are shared. Given the weak institutional capacity of subnational governments in many developing countries, it seems unlikely that they will have a strong hand in negotiating such contracts. The Indian High Powered Expert Committee for Estimating the Investment Requirements for Urban Infrastructure (2011, 101) puts it well. “Weak governments cannot rely on private agents to overcome their weaknesses nor can they expect to make the best possible bargains for the public they represent.”

4. Two Special Issues: Metropolitan Areas and Regional Equity

In many countries, specific problems call for special consideration, and perhaps for a different approach to infrastructure delivery. We discuss two such problems: how best to organize infrastructure provision in metropolitan areas and in their polar opposite—remote and sparsely populated rural areas—and how to take account of concerns about “regional equity” in the allocation of infrastructure funding and projects.

4.1 The need for asymmetry

Some policy-makers in developing countries would argue that in principle, political institutions should be structured to treat all citizens as equally as possible:

22. For detailed exploration of the structuring of PPP arrangements, see Engel, Fischer, and Galetovic (2010). For a skeptical view of the range of opportunities to exploit such possibilities, see Menard (forthcoming).
they should be symmetrical. On the other hand, if economic institutions are to produce relatively equal (symmetrical) outcomes they must often be structured asymmetrically: they must recognize explicitly the very different conditions existing in different regions and localities. Big cities—metropolitan areas—are different from other urban areas and call for a different treatment as regards fiscal decentralization. At the other extreme, small, remote municipalities in sparsely populated regions also call for a special set of financing and service delivery arrangements.

**Metropolitan areas**

The expected rate of migration to urban areas over the next two decades will strain the present network of urban infrastructure and generate new demands that will be difficult to absorb. The number of megacities (those with a population of more than 10 million) is projected to increase from 19 now to 27 in 2025, when about 10 percent of the world's urban population will reside in these cities. Of the projected 27 megacities, 21 will be in less developed countries. By 2025, there will be 48 cities with populations between 5 and 10 million, and three-quarters of these will be in developing countries (United Nations 2008).

Ingram, Liu, and Brandt (2013) estimate that annual urban infrastructure costs will be equivalent to about 3 percent of GDP for new infrastructure and 2 percent for maintenance. The estimates for India are that to meet projected needs, urban infrastructure investments (excluding maintenance expenditures) must increase from current levels of 0.7 percent of GDP to 1.1 percent by 2032 (High Powered Expert Committee 2011). The policy questions that governments in developing countries must answer are: Which level of government will deliver this infrastructure? With what degree of autonomy? How will it be financed?

For large urban areas, heavy infrastructure demands will come from three directions:

- First, natural population growth and new migrants impose demands for expanding, repairing, and modernizing capital facilities to accommodate the larger population.

- Second, industry and businesses, especially in the high-technology sector, demand infrastructure that will enable them to compete in the international marketplace. Transportation and information technology support feature prominently in their investment wish list.

- Third, it is expected that there will be about 2 billion slum dwellers in urban areas by 2050 (United Nations 2008) and services to provide a basic standard of living for these urban residents will be needed.

23. For a good discussion of the symmetry-asymmetry issue in the South African context, see Schroeder (2003).
Efficiency considerations might point toward decentralization in the provision of services in metropolitan areas. Governments in metropolitan areas have the advantage of knowing the needs and preferences of their population and businesses. But decentralizing expenditure responsibility to metropolitan areas is one thing; decentralizing to local governments within the metropolitan area may be quite another. The local units within the metropolitan areas may not be large enough to efficiently deliver infrastructure services. The underlying problem is the fragmentation of governance within large urban areas (Bahl 2013).

Local government structures in metropolitan areas typically emphasize one of three approaches to governance.

- A political fragmentation approach: governance is divided among numerous small municipalities, each operating with some degree of autonomy. This approach, which features home rule, is the one taken in the Mexico City metropolitan area (two states, a national capital district, and over 50 municipal-type local governments), and the Sao Paulo metropolitan area (39 municipal governments operating with relatively little required coordination).

- A functional fragmentation approach: the municipal government structure is overlaid by some special districts or public enterprises that provide infrastructure services to the entire metropolitan area. The emphasis here is on technical efficiency in the delivery of infrastructure services. Such an approach is taken in Mumbai, with state-owned parastatals serving as the metropolitan-area enterprises.

- A metropolitan government approach: general services are provided by an area-wide government with little by way of local governments. The emphasis of this approach is on coordination in the delivery of services in the area. This would describe the approach taken in South African metropolitan cities.

In the case of a politically fragmented government structure, the devolution of responsibility for infrastructure services does not lead to efficient outcomes. Jurisdictions are often too small to capture the economies of scale necessary for production efficiency and management skills may be limited. Further, the more fragmented the structure of government, the greater the problems with spillover effects from local investments in infrastructure. Finally, local governments in fragmented metropolitan areas seldom have broad-based taxing powers that would provide the basis for supporting a large debt issue.

If there is to be assignment of responsibilities for infrastructure services to subnational governments in urban areas, and if there are to be broad-based taxes in the system, the better approach would be either a metropolitan government or public enterprises whose service district includes the entire metropolitan area. On the other hand, within this general framework there may be ample scope for “neighbourhood” (or small municipality) works to be developed (and largely...
financed) by user-based groups through schemes capturing the value increments generated by such projects.

Remote localities
Rural local governments also call for different treatment. They are small, often remote, and have very limited resources. The literacy rate may be low, financial management capacity of the local governments may be limited, and there may be little capacity to design or deliver anything beyond the most basic infrastructure project. Own-source revenues may be all but non-existent because of the weak tax base, and user charges well below cost-recovery levels because of residents' poverty.

China is one of the most rapidly urbanizing countries in the world. Despite the rapid out-migration of recent decades, however, more than 600 million people still live in about 700,000 rural villages, many of which are in remote and sparsely populated areas. For these villages (and the townships in which they are situated) to become more viable, accountable, and at least moderately effective in providing basic public services, significant efforts and reforms are needed. Similarly, in India, another huge country with a large rural population, experience to date suggests that neither centralized provision of rural services nor more decentralized provision has worked very well. While generalizations are dangerous in either of these huge and heterogeneous countries, rural service delivery remains a problem and a very different one from that in the metropolitan areas.

One part of the solution to both problems is similar, however. In metropolitan areas, some kind of region-wide authority is needed to deal with area-wide infrastructure problems. In rural areas also, some kind of higher-level “regional” authority is also often needed, though not because of spillovers, since there are few such problems in small, geographically separated localities. Rather, the need is for technical, administrative, and financial support to provide basic local services in remote, small, and usually very poor communities. For example, small remote communities are unlikely to have an adequate tax base to be economically viable, to have adequate administrative capacity, to finance major capital expenditures, or even to take advantage of economies of scale (Kitchen and Slack 2006).

In these circumstances, infrastructure investments required to provide basic local public services must either be provided directly by a higher-level government or at least financed (with perhaps some limited in-kind or other local cost-sharing) by such a regional structure, whether a county, a second-tier municipality, or some kind of special district.

4.2 Regional equity
In the devolution of expenditure responsibility for infrastructure services, efficiency is not the only consideration. Fiscal disparities among regions in developing countries can be quite large, and the quality of infrastructure is usually

24. In some instances, the economically most efficient course of action might be to move people to some other area, but this is seldom politically possible and, at least in the case of border regions, not even politically desirable.
most wanting in the poorest places. A national infrastructure decentralization policy will, if anything, widen fiscal disparities because of the greater revenue mobilization capacity of richer jurisdictions and their superior capacity to deliver services. In addition, as Albalate, Bel, and Fageda (2012) show in detail for Spain, and as many “provincials” suspect is true in countries everywhere, centralized investment programs (e.g., in network projects) tend to be heavily related to proximity to the political capital city. Although infrastructure investment is more an instrument of centralization than of regional redistribution policy, it is nonetheless also commonly a vehicle for “regional policy” in the sense of providing “something for everyone” and perhaps even more for those who have less.

There are two main views about decentralized infrastructure responsibility in this context. One is that the ability to choose capital projects and to deliver them is just as important in small communities as it is in large urban areas. Devolution can work. There is evidence that in some cases, small local governments can produce small (local) public works projects at a lower cost and design them so that the probability of sustainable operations is relatively high. The problem is financing. In this case, decentralization with financing by intergovernmental transfers for smaller projects and vertical programs for the delivery of larger projects may be a viable strategy.

On the other hand, much of the evidence on successful devolution of infrastructure is about small settlements, including rural ones. There are questions about the how much of this success can be transferred to larger settlements or even to the rural sector more broadly. If not, infrastructure disparities may widen under devolution. In this case, the right strategy may be an asymmetric system in which larger urban places have heavy responsibility for delivering and financing their infrastructure, with infrastructure in smaller, less dense, and often poorer localities being delivered primarily by higher-level governments.

5. Conclusions and Guidelines for Policy
Theory tells us that an appropriately structured fiscal decentralization can lead to a higher quality of infrastructure services than will a fully centralized system. If the assignment of expenditure responsibility is correct and if the capacity to deliver services is in place, consumer-voters will be empowered to express their preferences for infrastructure services, services will be delivered with appropriate technologies and at lower costs, and there will be more willingness to pay and a better record of maintenance of public facilities.

Following the theory can indeed lead to some of these outcomes. But there is also plenty of evidence that a decentralized infrastructure structure fails the quality test. The problem is that the theory is based on a set of assumptions that may not apply to developing countries, and that successful infrastructure decentralization requires changes in the supporting governance and finance system that most developing countries are not willing (or able) to make. However, from this theory,

25. See, for example, the review of the evidence in Peterson and Muzzini (2005).
and from the practice in developing countries, we might draw out some policy rules to support successful decentralization of appropriate infrastructure services. This review of the theory and the practice suggests six policy reform areas issues that can lead to more success in the delivery of some infrastructure services by local governments.

First, there is the need for comprehensive infrastructure policy reform. Reformers should take a broad view of infrastructure decentralization and should consider and fit together all the components necessary to make the system work. The argument for local control depends not only on beneficiaries being local, but also on local control being more likely to meet local needs and, ideally, on local users’ ability and willingness to pay the costs. Only when local governance institutions ensure in an accountable way that those who benefit and pay are also those who decide what is done can all aspects of infrastructure decisions be assigned to local governments. A corollary is the importance of getting the structure right in order to benefit from management improvements. Improved procurement and financial management systems and the like will not achieve their goals if the responsibility for capital expenditure is assigned to the wrong level of government.

Second, a clear and appropriate assignment of expenditure responsibility for infrastructure is required. Public-sector functions must be sufficiently unbundled to ensure that they are assigned to the correct levels of government: both who is responsible for what and who is accountable to whom for what must be clearly delineated. While unbundling can be a decided improvement in that it recognizes the comparative advantages of each level of government in infrastructure service delivery, it can lead to complication and confusion and must be reviewed regularly. Clarity in assignment must be matched by accountability, in terms of both 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.

Third, local governments must be adequately accountable in political, administrative, and financial terms both to those whom they are supposed to serve—their residents—as well as to those above them in the governmental hierarchy who may be responsible for developing policy, regulating how it is carried out, and often financing the activities of lower-level governments. Such “dual accountability” is neither easy to design nor to implement, and requires fiscal planners to walk a fine line. On the one hand, unbundling schemes (such as allowing higher-level governments to set standards and require creditable compliance tests) call for upward accountability by local governments. So do conditional grants. On the other, too much expenditure-mandating by higher-level governments can take away the local autonomy that is the key advantage of fiscal decentralization of infrastructure. Ultimately, the local government must be accountable to its voting constituency. Mandates and conditions should be held to the test of accommodating externalities.

Fourth, the revenue mobilization efforts of local governments in most developing countries needs to be increased if local governments are to take on
more responsibility for delivering infrastructure services. In particular, the larger urban local governments need to be given access to stronger tax bases and encouragement to reform the tax base to which they already have access (namely the property tax). User charges must be ratcheted up significantly so that they approach cost-recovery levels. Arguably, the major impediment to increased revenue mobilization by local governments is the absence of incentives that are sufficiently large to offset the aversion of local politicians to tax increases. Here the reform agenda might include closing off the option of ad hoc capital transfers to cover infrastructure financing by local governments, the provision of incentives for increased local effort, and conditional grants that combine clearly specified conditions and local matches.

Fifth, borrowing and public-private partnerships have great potential as methods of financing capital projects, and in the latter case as a managerial approach to operations. However, governments in developing countries must be more realistic about the prospects of reaching either of these sources for efficiently financing local government infrastructure. Borrowing, even under a reasonable regulatory framework, can be part of the long-term infrastructure finance plan for local governments, but only if this is supported by a level of local revenue mobilization and transfers that will make principal and interest repayment as well as maintenance viable. Public-private partnerships are even more problematic for developing countries. For revenue-generating projects for which the user-charge phobia of local governments is not so great (e.g., telecoms, parking, etc.), private capital can be drawn. But for the high-priority services such as water and sewer, private investors will probably not buy in until they see a reasonable prospect for cost recovery and for operations that are not hamstrung by the regulatory framework.

Sixth, there is a strong case for an asymmetric approach to decentralizing infrastructure delivery and finance. The larger urban areas have a greater capacity to plan, deliver, and finance infrastructure services than do smaller and more rural local governments. The regime for infrastructure decentralization should recognize these differences, perhaps by limiting local government participation until local governments grow into the responsibility. Such a guideline would lead to more responsibility for delivering infrastructure in larger urban areas, as well as more responsibility for financing infrastructure with own-source revenues. Rural local governments will continue to rely more heavily on vertical programs of higher-level governments. This reliance may harm the equalization objectives of government, and it may slow down the capacity development of rural local governments, but it also may increase the productivity of total infrastructure spending.
Bibliography


Herstein, Lesley. 2012. *Adding value: Recognizing the link between engineers and municipal finance and governance.* Presentation at Institute on Municipal Finance and Governance, Munk School of Global Affairs, University of Toronto, May.


