Decentralization:
A Few Principles from the Theory of Fiscal Federalism

Bernard DAFFLON and Thierry MADIÈS

1 Revised version, June 2011
The Authors

Bernard Dafflon is a professor of public finance and public policies at the Department of Political Economy at the University of Fribourg.

Thierry Madiès is a professor of regional and international economics in the same Department. From 2008 to 2010, he was also a member of the Prime Minister of France’s Economic Analysis Council.

Acknowledgements

The authors would like to thank Réjane Hugounenq for her valuable comments. The authors are responsible for any remaining errors.

Foreword

The production of this addition to the Notes and Documents collection is part of a project launched by the AFD’s Research Department to study the territorial organization of state action and the subjects of decentralization/centralization. The aim is, by utilizing key theories and concepts from various disciplines (local public economy, political science, geography), to identify the primary determining factors behind the territorial organization of state action so as to establish the principles and modalities for empirical research.

This first paper uses the theory of fiscal federalism to analyze how state action is organized in countries from the standpoint of public finance. The elements examined here are, therefore, primarily economic in nature and relevant to consider in any centralization or decentralization process. This analytic paper will be completed by a series of working papers that will either clarify one of the issues raised in this document (assignment of functions, redistribution, etc.) or propose a contextual examination of the analysis using country studies.

Contact: Réjane Hugounenq:
Contents

1. Decentralization: attempted definitions .................................................................

2. What advantages are there to decentralizing? ......................................................

3. Assignment of powers to different levels of government ........................................
   3.1 Lessons from the traditional theory of fiscal federalism .................................
   3.2 Expanding traditional theory to make it more operational ............................

4. What type of taxation at what level of government? ..............................................

5. What autonomy? .........................................................................................................
   5.1 Financial autonomy ............................................................................................
   5.2 Budget autonomy ..............................................................................................

6. Budget transfers ........................................................................................................
   6.1 Justifications for budget transfers ....................................................................
   6.2 Criteria definitions ............................................................................................
   6.3 General principles ............................................................................................

7. Budget discipline and constraint ............................................................................
   7.1 Local governments’ borrowing capacity ............................................................
   7.2 Budget constraints ............................................................................................
   7.3 Assignment of “rights” to deficit and debt to levels of government ..............

8. Conclusion ..................................................................................................................

Bibliography
Decentralization: Attempted Definitions

During the past two decades, decentralization issues have generated keen interest in a large number of countries. This notion is, however, very difficult to define as it refers to a wide range of institutional arrangements. What is more, it is interesting to note that the circles of definition—and perhaps confusion—between decentralization and fiscal federalism usually overlap. Box 1 attempts to clarify these concepts. In addition, the motivations guiding the decentralization process differ from country to country. Thus, Bird and Ebel (2007: 5) introduce a series of studies on fiscal fragmentation by raising two fundamental questions from the start: “What is decentralization?” and “Why is decentralization occurring?” Ebel and Yilmaz (2002: 157) note on this subject that “developing countries are turning to decentralization to escape from the traps of ineffective and inefficient governance, macroeconomic instability, and inadequate economic growth,” whereas in post-communist countries, public-sector decentralization accompanies the movement toward a market economy and democracy. In Latin American countries, political pressure for more democracy also explains the decentralization movement. In Africa, a certain number of attempts at decentralization grew from a desire to preserve national unity by taking into account tribal and local particularities.

In Ethiopia, for instance, the decentralization process is a response to desire of maintaining national unity in an ethnically fragmented country. This is a unique experiment since other countries have, on the contrary, attempted to preserve their unity by maintaining a high degree of centralization, very often using coercion to force its acceptance. This raises the more general issue of institutions and the relations between the center and local governments. In the former French colonies, local governments are, in a sense, more agencies of the central authority than elected governments whereas, to a certain extent, the former British colonies allow greater room for autonomy.

The examples cited above could lead one to think that decentralization is the result of a more or less unilateral decision by the central government (even though the decision may be forced by the imperatives of preserving national unity, for example). In reality, there are two ways to think of decentralization: according to either a top-down or a

---

2 This is the case in two recent books: Bird and Ebel’s collective work (2007) uses the expression ‘decentralized countries’ in the title but refers in the chapters to ‘federal industrial countries’ and ‘new federal countries’; for their part, Ahmad and Brosio (2006) gave their collective work the title of Handbook of Fiscal Federalism, while its first chapter opens with ‘the political economy of decentralization’.

3 Does economists’ jargon not speak of decentralization when designating consumption and production decisions made by rational economic agents in response to a market price system?
bottom-up model. The objectives and implementation modalities are different in each model, as Table 1 shows.

### Box 1  Federalism and Decentralization: Two Distinct Notions

Before tackling the study of fiscal federalism and decentralization in its own right, it is necessary to clarify terminology so as to avoid all confusion between these two terms. They are neither synonyms nor antonyms. Their relationship can be clarified by three different approaches:

(i) From the standpoint of constitutional law, there are three possible distributions of state sovereignty. Thus, one distinguishes between unitary, federal and confederate government units. The constitution of a unitary state stipulates that the central government holds all sovereignty and that it may cede a greater or smaller share of this sovereignty to constituent units (which one designates by the generic—but imprecise—term of “decentralization”). In reality, the perfectly unitary state does not exist. A federation is a “system of voluntary self-rule and shared rule” (Kincaid, 2002: 7), in which the central government and the governments of the constituent territories decide by common accord how to divide powers and maintain their independence within their own spheres of authority. This independence is anchored in a federal constitution. A confederation is the union of several states that joint together around a minimal platform of collaboration while maintaining their own sovereignty.

(ii) “[I]n economic terms, most if not all systems are federal,” claims Wallace E. Oates (1972: 18), an economist of renown in this field. According to him, the difference between government systems is one of degree rather than category. At one end of the continuum, one finds unitary regimes in which all decisions are made by the central authority. At the other end, one finds the anarchist state. Between these two extremes, one finds more or less decentralized federal systems. Schema 1 portrays this point of view in a simplified manner.

(iii) Since the logics of centralization/decentralization and unitary/federal do not overlap, one could also present government systems by positioning them more precisely in a two-dimensional frame of reference, as shown in Schema 2. This would allow one to take into account the possible combinations, distinguishing between centralized federal countries (e.g. Germany) and “decentralized unitary countries” (e.g. Spain). Finally, the classification introduced by Lijphart (1999: 189) is a compromise between the continuum and the quadrant insomuch as it distinguishes between decentralized federal countries, centralized federal countries, semi-federal countries, decentralized unitary countries and centralized unitary countries.

Let us specify that we use the expressions “sub-central levels of government” (in OECD terminology) and “local governments” indiscriminately to designate both the intermediary level of government (states, regions, cantons, provinces) and local government units. This can be confusing since, in federal countries, the intermediary level has substantial autonomy whereas in unitary countries such as France, for instance, the regional level is viewed as a local government in the same right as the departments or communes.
Let us note that the “unitary/federal” dimension is viewed from the constitutional standpoint: it is therefore a country’s basic status that determines its position along the horizontal axis in the schema. The “centralization/decentralization” dimension, however, comes from a concept that is much more difficult to grasp because it depends above all on a quantitative and qualitative measurement of the assignment of powers to the various levels of government. The OECD has been working to refine this reference concept for several years (Blöchliger and King, 2006). Among other things, some believe that the generic term “decentralization” is not very appropriate because it is too vague, and that one should distinguish between deconcentration, delegation and devolution.

The first type of decentralization (which is the one to which we shall refer almost exclusively in what follows) is top-down in nature and gives the central government’s preferences a “strategic” advantage. It is usually implemented in unitary systems of government in which the central government always retains ultimate control of arbitration powers (Table 1). This decentralization process implies a transfer of power and responsibilities previously assigned to the central government to specialized entities (“functional” decentralization—in France, these entities are incorporated public establishments whose jurisdiction is determined by their articles of association) or to entities whose action lies within a given territorial map (“territorial decentralization”, to use an unofficial expression). This process can also be implemented in federal governments in which the central government shares government sovereignty with the constituent states but the latter do not dominate the central level (see Box 2).
Much rarer in practice, the second type of decentralization (bottom-up) concerns federal governments in which priority is given to local preferences, and local governments have substantial autonomy. There are generally two types of questions: (i) the question of which functions must be transferred to a higher level of government (against the backdrop of the subsidiarity principle); and (ii) the question of competition and/or cooperation between local governments and between these local governments and the higher level of government.

Tables 1 and 2 make it possible to (i) compare the underlying logics of bottom-up and top-down decentralization, and (ii) address decentralization or fiscal federalism from the standpoint of public finance.

Table 1 Two Types of Decentralization

<table>
<thead>
<tr>
<th>Forms of (De)Centralization</th>
<th>Top-Down = become local</th>
<th>Bottom-Up = remain local</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Objectives</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• shift fiscal constraints to the local level (transfer of costs rather than devolution of powers)</td>
<td></td>
<td>increase local-level innovation potential</td>
</tr>
<tr>
<td>• increase national wellbeing (better match of supply to demand) by maintaining the minimal standards of available decentralized public policies (supervisory standards)</td>
<td></td>
<td>maintain the diversity of “local public goods (LPGs)/local financing (taxes or user charges)” combinations allowing one to “vote with one’s feet”</td>
</tr>
<tr>
<td>• increase allocative efficiency (better adapt the supply of public goods to local preferences)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• ensure good “governance” (equivalence of Olson’s three spheres; fiscal responsibility or accountability)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Assessment Criteria</strong></td>
<td>contribution to objectives set by the center</td>
<td>protection of local governments from the center and maximization of local decisions</td>
</tr>
<tr>
<td><strong>Modalities</strong></td>
<td>deconcentration</td>
<td>local autonomy</td>
</tr>
<tr>
<td>• delegation</td>
<td>subsidiarity</td>
<td></td>
</tr>
<tr>
<td>• devolution</td>
<td>decentralized cooperation</td>
<td></td>
</tr>
<tr>
<td>• vertical relations between the center and the regions (center and communes, respectively regions and communes) dominate</td>
<td>horizontal and vertical competition</td>
<td></td>
</tr>
<tr>
<td>• the center’s preferences dominate; the principal (center) – agent (the regions or communes) model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• information gap favoring the center</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Dominant Model</strong></td>
<td>central/local conflict (if society is heterogeneous, local decisions differ from those preferred by the center)</td>
<td>local autonomy dominates</td>
</tr>
<tr>
<td>• local autonomy dominates</td>
<td>subsidiarity</td>
<td></td>
</tr>
<tr>
<td>• cooperation and competition</td>
<td>local preferences dominate</td>
<td></td>
</tr>
</tbody>
</table>
Top-down ("territorial") decentralization calls on different notions that are often intertwined in reality. Traditionally, a distinction is made between the notions of deconcentration, delegation and devolution (Gauthier and Vaillancourt, 2002).

- **Deconcentration** consists of giving power over certain decisions to agents that exercise their functions in a given territorial district while remaining under the hierarchical authority of the central government. Lower levels of government (which can be simple administrative relays) are, in this case, subordinate in fact to the central authority. Deconcentration generally aims to improve the operational efficiency of the central government’s action within a given administrative district (let us not forget the words of a famous professor of French law who said, “deconcentration is the same hammer hitting you on the head, but with a shorter handle”).

- **Delegation** (sometimes also called an agency relationship) designates the transfer of power and responsibility from the central government to often semi-autonomous entities within a clearly defined field of responsibility (in the framework of “functional” decentralization, let us remember that these entities have independent legal status). These entities with independent legal status and autonomous budgets are generally obliged to report to the central government. Local governments generally provide certain very specific services that are delegated to them by the central government or the next-higher level of government (in this case they are the recipients of delegated powers and resources).

- Finally, **devolution** is the most advanced form of decentralization in that it transfers powers and responsibilities to public bodies elected by their constituencies ("territorial” decentralization in France refers only to this meaning). Responsibilities and resources are transferred to local powers (often local governments) that have considerable autonomy of decision as to how to use these resources in their area(s) of responsibility and in a legally acknowledged administrative jurisdiction. Devolution goes hand-in-hand with political decentralization when the local powers must account for their decisions before elected assemblies. The effect of this is to favor grassroots democracy. Devolution is habitually confused with fiscal decentralization when it comes with a clear assignment of financial relationships and powers to levels of government that have the ability to raise financial resources (fiscal autonomy) and have substantial budget autonomy (in regard to spending). It should be emphasized that, in practice, the term “decentralization” is often used to mean devolution alone, which can cause confusion. It should also be noted that, in most countries (even federations), these three methods can exist simultaneously and complement each other. For instance, in Switzerland, the cantons have kept a large amount of latitude in the field of education (which would equal a devolution of the top-down type of decentralization), but their environmental policies are heavily determined by federal standards (which would equal delegation, with the communes executing the function in compliance with federal standards, nothing more).
### Table 2  Public Finance According to the Form of Government

<table>
<thead>
<tr>
<th>System of Government</th>
<th>Decentralized Finance</th>
<th>Fiscal Federalism</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>centralized federal (Germany)</td>
<td>decentralized federal (Switzerland)</td>
</tr>
<tr>
<td></td>
<td>decentralized unitary (Spain)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>centralized unitary (France)</td>
<td></td>
</tr>
</tbody>
</table>

#### Constitutional Framework (institutions)
- Institutions are organized in a normative “constitution” (number of levels, territorial map, powers, financing, and resource transfer rules).
- The constitution establishes the local governments and sets the rules on inter-governmental relations.
- All powers not explicitly assigned to the local governments continue to belong to the center.
- The center’s preferences are dominant.
- The territorial distribution dates from before the constitution, set ex post. The constitution ratifies the limits of the local governments.
- The constitution is a reflection of a voluntary association of sovereign local governments (a confederation), and sets the rules on inter-governmental cooperation.
- Powers belong in fact to the local and canton governments. The constitutions (canton, federal) must explicitly mention the powers transferred upward. The same holds for sources of financing (except transfers).
- The constitution = all the rules + consensus requiring a double majority (voters and cantons) for upward transfers of powers and taxes.

#### Types of Inter-Governmental Transfers
- importance of specific incentive transfers
- redistributive transfers targeting equality of individual situations in regard to LPGs and taxes
- respect for autonomy of decision: few specific incentive transfers
- redistributive transfers aiming to limit disparities but without regard to individual positions

#### Inter-Governmental Relationship Model
- typical “principal – agent” model
- in centralized unitary countries, the principal defines financing powers and rules in function of its own objectives, information gaps and moral hazards
- “bargaining among principals” model

#### Typical Cases
- centralized federal countries (Venezuela, Austria, India)
- decentralized semi-federal or unitary countries (Spain, Netherlands, Sweden)
- centralized unitary countries (France, United Kingdom)
- but also: “impure federalism” (Germany, USA)
- “dialogue/diplomacy” between provinces and the federal government in Canada
- Switzerland: bodies of cooperation between the Confederation and the cantons; “canton conferences” for functional coordination; consultation procedure; initiative and referendum rights
Box 2  What Type of Federalism Are We Discussing?

The notion of “federal fiscal constitution” designates the rules of the game or standards—formal, institutional or other—in principle inscribed in a country’s fundamental charter, which governs public policy choices and spending decisions. The federal constitution is not the same in Canada, Australia or Switzerland. It determines which of the various levels of government does what in the field of public spending and who obtains what in the field of tax levies. In Switzerland, the “fiscal” part of the federal constitution reflects the so-called subsidiarity rule under which the cantons manage, in principle, all economic areas that they have not ceded to a higher political and administrative level.

*Dual federalism* is characterized by a clear separation of powers between the levels of government. In other words, the powers attributed to each level are exclusive (layer cake federalism). Attention: this does not mean that the full range of powers is distributed in a balanced fashion between the federal level and the regional/provincial or local level. It simply means that when a power is assigned to a level, sovereignty is, in principle, not shared. Dual federalism is very rare in practice.

*Cooperative federalism* is characterized by a greater interdependency between the levels of government. Decisions are made after a relatively complex consultation process among territorial actors or local governments: one can see a close and constant decision-making exchange between “regions” and the “center”. In this case, there is a “vertical fragmentation” of powers with each level of government contributing to public policy (marble cake federalism).

*Competitive federalism* is characterized by competition between (local or regional) governments in regard to public policy. In order to increase their tax base, local governments offer bundles of public goods that are attractive for taxpayers (individuals and companies) and attempt to minimize their tax rates. In fiscal federalist thought, the term “competition” often has connotations of waste, overlapping powers, and inefficiency. Other approaches defend competition’s contributions to reducing the state’s share (taming the Leviathan) and promoting public policy innovations (laboratory federalism).

*Federalism of execution* is characterized by the fact that the federal government keeps its responsibilities and powers for certain specific functions but has regional or local governments execute them. The “principal” in the “principal – agent” relationship is the central/federal government that decides and dictates the rules to follow for a given function for which it is fully accountable. The “agent” is either a decentralized government (the cantons, for example) when speaking of federalism of execution or delegation, or a decentralized office of the central government, one of the center’s “regional agencies”, when speaking of deconcentration. Naturally, the same distinction can be reproduced on the regional/local level.

In Switzerland, the field of environmental protection is a good example of federalism of execution. Environmental policy targets and the environmental standards to apply or attain are set by federal law. The cantons are in charge of coordinating implementation within their respective spheres. The communes’ job is to apply and execute measures and finance them according to the polluter pays principle. At the start of a project, incentive grants (see Chapter 6) assist execution.

Table 2 also examines the logic of decentralization and fiscal federalism, but does so from the constitutional standpoint of public finance. The table is read in relation to the vertical axis of Schema 2 in reference to systems of government. In general, economic analysis of the constitution allows one to distinguish between two approaches.
(i) In the case of “decentralized public finance”, all powers and resources belong *a priori* to the central government. The text of the constitution must therefore explicitly delegate or devolve powers or tax categories to the regional or local governments so that they have the right to act within the prescribed limits (moving from the bottom up on the vertical, financial axis of Schema 2, from the left).

(ii) Inversely, in its pure form, fiscal federalism assigns powers and resources *de facto* to the local level (the top of the vertical, fiscal axis in Schema 2). A constitutional procedure must therefore be triggered to transfer one or the other to the regional, then central, level (centralizing tendency, from the right).

Thus, in Switzerland, transferring a power or tax from the local level to the canton level normally requires a simple majority vote modifying the canton’s constitution, whereas transferring powers from cantons to the Confederation or the inverse demands a modification of the federal constitution, which in turn requires a double majority of voters and cantons. Vertical fiscal relations are more limited in fiscal federalism, first out of respect for the autonomy of the sub-central units, and because of their responsibility for assuming the fiscal consequences of spending decisions. But also because the target of redistribution is often different: whereas, in decentralized finance, redistribution aims to allow residents access to comparable services without a heavier financial burden, in fiscal federalism the aim is, rather, to bring disparities between local governments down to a politically acceptable level without reference to individual positions. This (Swiss) reference could lead one to believe that it is a sufficiently pure form of federalism to place it unequivocally in the upper right triangle of Schema 2. This is not the case. Seen in budgetary terms, fiscal federalism itself can have several contents. Box 2 attempts to summarize these concepts.

Now that definitions have been attempted, the following points aim to provide a partial guide to interpreting a certain number of issues pertaining to the advantages of decentralization (Chapter 2), the assignment of powers (Chapter 3) and fiscal resources (Chapter 4) to levels of government, measuring fiscal and budget autonomy (Chapter 5), transfers between levels of government (Chapter 6), and finally the nature of “hard” or “soft” budget constraint (Chapter 7).
2

What Advantages Are There to Decentralizing?

The advantages attributed to decentralization by the fiscal federalism theory can be judged in light of Musgrave’s (1959) analysis of the major state functions. A distinction is generally drawn between the three types of functions fulfilled by the public sector (Musgrave’s famous “trilogy”): resource allocation, redistribution, and stabilization of economic activity. Included in the modern version of this last function are the state’s macroeconomic functions.

- Traditional thought is that the stabilization function must remain the central or federal government’s responsibility. One of the arguments advanced in favor of centralizing the stabilization function, which we shall discuss in greater detail shortly, is based on the idea that the tax revenues of sub-central local governments must be very stable and must thereby come from tax bases that are little elastic and little mobile. This in turn obviously restricts their ability to play a role in stabilization. (Readers will note that this argument is of the same nature as “the chicken or the egg.”) The other pro-centralization arguments are that: (i) that local economies are small open economies and, as such, there is little chance that stabilization efforts initiated at local level will bear fruit in this same local arena (Balassone and Franco, 1999); (ii) inversely, local government units could adopt a waiting strategy, letting the other local governments implement stabilization policies whose effects would spill over in their favor; and, finally, (iii) limiting local budget deficits and local public debt raises problems—with positions as divergent as they are nuanced, known under the terms of budget “discipline” versus “accountability”, as we shall see in Chapter 7—because, in this area, financial markets do not function properly to attain budget discipline (Ter-Minassian and Craig, 1997:162).

- In regard to redistribution policy, the mobility of individuals (or even economic activities) is what constitutes the bulk of the argument for centralizing (interpersonal) redistribution policies. It is claimed that the wealthiest individuals will try to avoid a higher tax burden by “voting with their feet”, whereas poor individuals will move to areas where the local governments are most generous with their social transfers. One will note the positions taken by Ahmad, Hewitt and Ruggiero (1997) who recommend redistribution in kind through standardized services and identical access to certain services, classically education, sanitary services and basic health care.

---


5 Let us emphasize, however, that centralizing the redistribution function can be accompanied by a deconcentration of social spending or a delegation of the same to elected local governments with a very strict agency mandate. In this case, social benefits will be more or less equally generous throughout the national territory.
The mobility argument seems, however, somewhat inoperative in the case of developing countries insomuch as the wealthiest individuals have, in reality, a very limited choice of places to live where they can take advantage of amenities that correspond to their standard of living, which restricts their real mobility. Similarly, poor people from rural areas migrate to the periphery of urban centers for reasons usually linked to the search for jobs or a brutal worsening of their economic conditions, which leaves little latitude to deliberately decide to take advantage of more generous social services. On this last point, we will, however, note that there is nothing to indicate that the poor and the wealthy in developing countries would not move to receive more services (the poor) or avoid overly high taxes (the wealthy) if these countries’ taxation and redistribution systems were analogous to those in developed countries.

- The question of assigning the allocation function (what must be produced? for whom? at what “fiscal cost”?) is, in reality, the most important question in our opinion and the one on which we shall concentrate. The underlying idea is that devolution (when it confers consequential autonomy on local governments) best makes it possible to match local public services to citizens’ preferences (and, to a certain extent, also to companies’ needs). Yet, for this—and it is far from guaranteed in developing countries—devolution must be accompanied by a democratic participation process with a solid institutional foundation. By bringing the policy-making process closer to citizens, decentralization gives them greater control over the production cost of the services offered. The allocative and productive efficiency of local public goods (LPGs) is therefore improved. However, this “three-P trilogy”—preference, participation, proximity—has at least two exceptions: economies of scale and spillover. For public goods and services with scale-based production, larger-scale production (supra-communal or even regional—therefore with a centripetal tendency) is necessary for average costs to be as low as possible (see below). In this case, balance must be found between lowering costs and the fact that the supply of such services will be “less of a match” to local preferences (unless preferences are homogenous). An analogous problem arises when there is spillover in the consumption of certain public goods and services between local governments (between communes, for example). In this case, economic theory tells us that the supply of local public goods and services will be sub-optimal and that a cooperative solution will be necessary. Oates (1972) summarizes these considerations in his famous decentralization theorem: “each public service should be provided by the jurisdiction having control over the minimum geographic area that would internalize benefits and costs of such provision.”

- Among other things, decentralization makes it possible to curb the discretionary power of elected officials because it comes with greater competition. This competition can take two forms: first, the mobility of individuals and economic activities between local government units and, second, the possibility for voters to compare the performance of their elected officials with the performances of elected officials in neighboring local governments (yardstick competition).

The question of the effects of budget competition — sometimes limited to the tax component alone — remains largely unanswered. Tiebout (1956) and his famous “vote
with their feet” hypothesis emphasizes that individuals chose their place of residence taking into account not only the taxes that they will have to pay but also the local public goods (services) that they can receive. This mechanism leads — through a set of restrictive, even desperate, theoretical hypotheses — to an economically efficient situation as the private market would. Tax competition (in this case only looking at the tax component) is, among other things, seen by some authors (Brennan and Buchanan, 1977, 1980) as a way of limiting government’s supposed predatory behaviors. But it can also turn out to be a source of economic inefficiency when it leads to “one-upmanship” in tax rate reductions and, ultimately, to a “prisoner’s dilemma” type situation. Indeed, it is easy to understand that if the tax base is little mobile or little elastic (that is to say, not very sensitive to tax rate variations), then a drop in rates leads to a drop in tax revenues, and therefore a worsening of the public services provided, for all competing local government units (because the increase in the taxable base does not offset the rate reduction). Among other things, as we shall see below, economic analysis shows that a local increase in fiscal competition can increase predatory fiscal behaviors when the local level shares its taxable base with a higher level of government!

- Finally, another virtue is attributed to decentralization: fostering experimentation insomuch as the good practices of certain local government units can be replicated and benefit other local government units. Obviously, some will insist that a central government can also “experiment” with economic policy measures by targeting certain parts of the country initially through, for instance, the deconcentrated authority that represents it at the local level (a governor or prefect nominated by the center). One crucial difference, however, is that it is in the interest of the deconcentrated authority in charge of applying these measures to overestimate the gap between the benefits (that are difficult to measure for public goods, even local ones) and the costs (because there is an information gap: the deconcentrated authority that produces and therefore controls production costs does not necessarily inform the principal of them). The deconcentrated authority can allow itself to behave in consequence because it is under the central government’s authority. A local government, for its part, does not have this strategic possibility: it is accountable to its own taxpayers.

Weingast (2006) emphasizes that the decentralization process must be designed as an experimental process that should be conducted in stages and not uniformly and simultaneously in the entire territory. The idea is first to chose (or, better yet, foster a bottom-up selection process) the regions or local governments in which resistance to change is weakest (local LPG preferences close to the desired allocative target, administrative capacity to implement decentralization, little resistance from local elites, etc.) in order to trigger a bandwagon process thanks to a demonstration or comparison effect on the other regions (is this not what China ultimately did with Guangdong Province?). Two problems can, however, be noted. (1) If the center selects regions or local governments where resistance to change is weakest, what system of decentralization should be recommended: delegation or devolution? Who decides what and how? (2) In practice, other local governments imitate “good experiments”
fairly quickly and, accordingly, experimentation in reality results in forms of budgetary mimicry, consequently narrowing the range of possible innovations.

**Box 3  Yardstick Competition as a Way to Limit Predatory Behavior Among Local Governments**

The interactions in regard to neighboring local government units’ fiscal and budgetary choices can be based on two types of mechanisms. The first is tied to the mobility of tax bases (see the abundant writings on tax competition) that leads local decision-makers to take into account tax and more generally budgetary variables (the quality of public infrastructures and certain local public services is also an important factor in decisions on the location of economic activities and the most mobile people) of neighboring local government units. The second is based on the characteristics of the local political market. This approach in the literature grew out of two hypotheses. The first hypothesis makes government decision-makers into potential Leviathans whose predatory behaviors can be regulated by the intermediary of the electoral system and the political sanctions that it can exact. This is the political voice hypothesis formulated notably by Hirschman (1970). The second hypothesis, from Salmon (1987) in particular, gives voters the ability to compare the performances of their elected officials and their neighbors’ and, thus, the ability to identify possible opportunistic behaviors (comparing political performances would, in a way, replace the “exit” of individuals when this is impossible). This is the yardstick competition hypothesis applied to analysis of political behavior. These two hypotheses make political life a subtle uncooperative game in which some try to exploit others while the latter seize every opportunity to detect predatory behavior. Indeed, whenever voters can see what goes on in neighboring local governments or the like, they can draw conclusions from this comparison. In this way, existing local government units must take into account not only the criticisms and promises of their direct competitors in the territory they run but also the performances of neighboring elected officials.

Two theoretical conclusions seem to emerge from this line of thought. First, elected officials will have a tendency to imitate each other to avoid being stigmatized. The result of fiscal or more general budgetary mimicry gives its name to this class of models. Secondly, this mimicry will result in tax rates that are higher than necessary to finance the level of public goods desired by voters, leaving elected officials with a rent. Empirically speaking, in nearly all studies (although they generally examine industrial countries), it appears that local governments’ fiscal and budgetary decisions are the result of mimicry in that they often depend heavily on the decisions made by neighboring local government units (Madiès, Paty, Rocaboy, 2005). This observation is very important because it places limitations (whether through tax competition or yardstick competition) on local government’s fiscal and budgetary choices.

- Traditional theories of fiscal federalism (Musgrave, 1959; Oates, 1972; Rubinsfeld, 1987) have, however, a very large normative content in that government decision-makers are generally assumed to be “benevolent” and maximize the social wellbeing function. Reality is, obviously, something else entirely and the behavior of elected officials and government decision-makers depends to a large extent on the institutions that define the framework for their action and the incentives they face (the term “institutions” should be interpreted in a broad sense). In short, decentralization is not good or bad in its own right—it depends on how it is

---

6 In the case of developing countries, one can raise specific question of whether there is a real alternative or simply parallel groups of predators (direct competitors are waiting to be elected to take their turn helping themselves).
implemented and the incentives it provides. “Decentralization is neither good nor bad for efficiency, equity, or macroeconomic stability: but rather that its effects depend on institution-specific design” (Litvak, Ahmad and Bird, 1998).

A certain number of advantages conferred by traditional theories of fiscal federalism must be amended if one wants to apply them to developing countries for several reasons. First, it is not at all certain that, in the poorest countries, individuals’ preferences are what steers the supply of local goods and services, as some prominent scholars of fiscal federalism assume. Indeed, in the poorest countries and others, it is likely that extreme poverty “homogenizes” individuals’ demand behaviors (“constrained demand”). The crucial question then becomes meeting the population’s vital needs (i.e. supplying basic goods and services). This can be done nationally or in a deconcentrated manner or even by forcing local governments to implement, as agencies, policies set by the center. A recent study in Colombia (Faguet and Sanchez, 2007) shows that the decentralization process resulted in a significant increase in vital public services in most of the small communes that had been neglected by the central government: the local authorities focused on essential services because that is where pressure from the population was felt. The real reasons for decentralization in developing countries can, thus, be found in the need to take local or tribal particularities into account. What is more, it appears that in many countries votes, when there are local votes, do not necessarily express citizens’ preferences. Sometimes, under threat, if the results are not in line with the center’s desires, the center may levy budgetary sanctions, notably by arbitrarily cutting financial resources in the local government units seen as opposing or “voting badly”.
3

Assignment of Powers to Different Levels of Government

There are at least three good reasons to state outright that allocating fiscal resources and budget transfers among levels of government must come after a clear assignment of spending powers.

- If the responsibilities of the various levels of government are poorly defined, then obviously the resources will also be inadequately defined in regard to both the amounts needed to finance spending and in regard to their nature (as we shall see, not all spending is intended to be financed with the same type of funds).

- Next, if responsibilities are not clearly or specifically defined, one can fear too much discretionary power by the various governments. This could lead them to “skew their budget choices” in function of strategic or partisan considerations to the detriment of longer-term (less visible) spending that would help improve growth in the regional economy. Hence, there is a need for the major founding principles behind decentralization to be embedded within a fundamental law and for the framework regulating relations between levels of government to be well defined.

- Finally, the accountability principle—one of the advantages of decentralized structures—can only be implemented with a clear assignment of powers. In the (unfortunately very frequent) case of overlapping powers among levels of government, some levels—notably those furthest from taxpayer-voters—can successfully engage in rent-seeking behaviors or profit from X-inefficiencies while “hiding” behind the levels closer to taxpayer-voters.⁷

3.1. Lessons from the Traditional Theory of Fiscal Federalism

We have already addressed the issue of assigning the public sector’s functions of stabilizing economic activity, redistributing revenue and allocating resources to the various levels of government. Here, emphasis is placed on a few principles to guide the assignment of certain functions traditionally fulfilled by the public sector to one (or more) decentralized level(s) of government. It is important, beforehand, to note that the institutional apportionment that determines the territory over which the political

⁷ X-inefficiencies refers to a situation in which the producer agents who know they are little or not at all supervised—or because they alone master the production function and do not have to inform the principal (information gap)—become more lax in meeting the demands of productive efficiency. They avail themselves of a production or administration comfort that is difficult to detect from the outside. These inefficiencies are labeled X because we do not know exactly who benefits from this situation (probably each agent in the production sequence).
power will exercise its sovereignty and the functional network of available public services are not the same. Based on this, two types of approaches can be taken. One solution is to try to define the optimal production and consumption area for a public service and possibly “create” local government units in charge of providing these services. For example, this is the approach taken by club theory and, with a different perspective, by functional overlapping competing jurisdictions or FOCJs (Eichenberger and Frey, 2007). A second, more pragmatic, solution that we prefer (unless one is a fan of “institutional mille-feuilles”) consists of studying how existing local governments can provide these services as efficiently as possible. Institutional combinations can, in this case, take very different forms; inter-commune cooperation is one possibility among others (Dafflon, 2000; Dafflon and Ruegg, 2003).

The two solutions are not an exhaustive response to the question of what criteria allow one to determine which degree of decentralization or centralization must prevail in the provision of this or that service. Oates (1972) advances five criteria (Table 3) by which to assign functions to levels of government. Like Oates’s model, and in the same spirit, most later models introduce an interpretation matrix containing criteria pointing to centralization compared with other criteria justifying decentralization. As these criteria have opposing forces, one must describe and measure their impacts in a precise manner, and then weight them. The result of this assessment and weighting decides the issue of (de)centralization. One will immediately see that, while economists have a key role to play in describing and measuring each economic criterion in the list below, weighting is a political matter.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Decentralization</th>
<th>Centralization</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. preferences</td>
<td>heterogeneous</td>
<td>homogenous</td>
</tr>
<tr>
<td>2. economies of scale</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>3. spillover</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>4. congestion</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>5. decision costs</td>
<td>if they increase with the size of the group</td>
<td>if they decrease with the size of the group</td>
</tr>
</tbody>
</table>

(C1) The first criterion deals with the degree of preference heterogeneity across the institutional territory. If individuals with similar preferences (for instance because their living conditions are specific) reside in a limited territory, then decentralization can be a way to best meet specific needs. Similarity of preferences is not easy to grasp but it must, for example, take linguistic particularities into account (this issue is important.

---

8 The economic dimension of public functions, notably to have economies of scale, is very often larger than the political territory of a single commune. Hence difficulties assigning powers to the local level and organizing production. Dafflon and Ruegg (2001: 15-17) observe that the functional territory (which refers to the optimal production dimension for a local public good or LPG) does not coincide with the institutional territory (local political boundaries), with both differing from the relational territory, which is often much larger (the “private” space in which economic and social activities of production, consumption and leisure take place).
when the goal is to determine within a function—in this case education services—whether school programs, one of the education “inputs”, should be defined at the central or local level. It should also be emphasized that if ethnic or linguistic communities are present in the same territory, then certain services can be delivered in the framework of the powers assigned to these communities. This is, however, very difficult when a linguistic minority is scattered throughout several areas of reference.

(C2) The second criterion is tied to the scope of technical economies of scale when producing the public service in question. These economies of scale can differ according to the function in question. This implies different functional networks for each of them. In practice, this operating logic can be found in communes’ single-vocation syndicates. The functional territory can be analyzed from the standpoint of production for an LPG if economies of scale justify production that goes beyond the commune and requires collaboration between communes or from the standpoint of consumption if a local government’s production of an LPG has spillover effects. Potential economies of scale are not an ipso facto justification for moving to a new functional, inter-commune scale. Public-private partnership (PPP) solutions can also internalize economies of scale. Communes continue to control supply individually, while they jointly externalize production to take advantage of higher yields.

(C3) The third criterion deals with the positive and negative territorial externalities of LPGs. There are two types of spillover: production and consumption.

- Production spillover happens when the LPG produced in local government unit (LGU) A has effects on other adjacent LGUs without the latter taking part in the decision or sharing the cost of the LPG. Example: If local government A takes air protection measures, the “clean air” thus gained does not stop at the border—it spills over and neighboring local governments benefit from a cleaner situation.
- Consumption spillover happens when the LPG produced by LGU A can be consumed by residents of adjacent LGUs who move to A to take advantage of the service without paying when it is not possible to exclude them. Example: the centrally-located communes that produce cultural LPGs that the inhabitants of outlying communes benefit from without participating in the deficit or public subsidies for these activities.

As is the case with economies of scale, the typical solution to positive spillover is to increase the institutional territory to adjust it to the functional territory in order to cause the circles of decision-makers, payers and beneficiaries to coincide. One alternative would be that the communes whose residents benefit from the spillover compensate the local government that produces the LPG causing the said effect. This solution brings the circle of beneficiaries and the circle of payers together.

(C4) The fourth criterion, the inverse of the preceding criterion, is that of negative territorial externalities, more specifically congestion. Commuting between LGUs allows

---

9 This solution is common in Switzerland for early childhood structures, family assistance and in-home care organizations, medico-social establishments for at-risk or dependent elderly people, and fields of technical production such as waste removal or public urban transit.
the commuters to benefit from the destination local government’s LPGs. If the quality of consumption of a public service depends on the number of consumers, an increase in consumers would, after a certain threshold, generate costly congestion. The reason that congestion generates a loss for the producing local government is that the commuters take into account the average costs that they personally pay, but not the marginal costs that they impose on all of the members of the group in which they arrive. An efficient solution requires either that one slow the arrival of new users (both permanent users and commuters), or that the destination group bill the new users for the cost differential, or that the group at the origin of the commute repay (subsidize) the destination group. Alternatively, extending the functional area of reference would make it possible to find adequate solutions (such as public transport to the center with parking lots in outlying areas).

(C5) In the model proposed by Oates, the last criterion is decision costs. For a given population, the optimal structure for the public sector is said to be a division in as many links as required by the LPG supply. Each series is said to guarantee a perfect match between the public service’s economic dimension and the size of the group that assumes budgetary responsibility for it, the functional network being in this case identical to the institutional network for each function. However, having users belonging to a multitude of groups (in fact: as many groups as LPGs consumed) would create high decision costs because they would have to participate in as many decision-making processes as required by how the provision of public services is organized. Among these costs, we shall mention: the election costs of each group’s executive bodies; the administrative cost of organizing and managing supply; the cost of informing the participants that would have to set budgets. One way to reduce these costs is to form multiple-purpose inter-commune syndicates providing a bundle of LPGs whose networks coincide sufficiently.

Obviously, these five criteria are not the only possible elements on the list. But the principle remains the same even when adding to the list: it is a matter of comparing the centripetal and centrifugal tensions to find the right balance for decentralization to be implemented. Thus, a frequently mentioned additional criterion deals with emulation and competition, which are associated with the existence of local governments (see the discussions below on the subjects of tax competition and yardstick competition).

Table 4 is an attempt to classify spending by level of government, using this logic. This classification of functions is rudimentary but it has the merit of insisting on the distinction to be made between regulatory functions and public service supply functions. In many cases, the chosen criteria—which are those of traditional fiscal federalism—do not by themselves make it possible to assign a function to a level of government. Under “comments”, the last column on the right gives other classification criteria that serve to enrich and extend the analysis to make it operational. Among other things, this classification tells us nothing about the main problems that arise at the local level because, by opting for the “state/region” or “local government” options, it masks in fact the question of collaboration between local governments (communes) that can take very diverse forms.
<table>
<thead>
<tr>
<th>Function</th>
<th>Policy, standards &amp; oversight</th>
<th>Provision/administration</th>
<th>Production/Distribution</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interregional &amp; International conflicts resolution</td>
<td>U</td>
<td>U</td>
<td>N,P</td>
<td>Benefits &amp; costs international in scope</td>
</tr>
<tr>
<td>External trade</td>
<td>U</td>
<td>U,N,S</td>
<td>P</td>
<td>&quot; &quot;</td>
</tr>
<tr>
<td>Telecommunications</td>
<td>U,N</td>
<td>P</td>
<td>P</td>
<td>National regulation not feasible</td>
</tr>
<tr>
<td>Financial Transactions</td>
<td>U,N</td>
<td>P</td>
<td>&quot; &quot;</td>
<td>&quot; &quot;</td>
</tr>
<tr>
<td>Foreign Direct Investment</td>
<td>N,L</td>
<td>L</td>
<td>P</td>
<td>local infrastructure is critical</td>
</tr>
<tr>
<td>Defense</td>
<td>N</td>
<td>N</td>
<td>N,P</td>
<td>Benefits &amp; costs national in scope</td>
</tr>
<tr>
<td>Foreign Affairs</td>
<td>N</td>
<td>N</td>
<td>&quot; &quot;</td>
<td>&quot; &quot;</td>
</tr>
<tr>
<td>Monetary policy, currency, banking</td>
<td>U, ICB</td>
<td>ICB</td>
<td>ICB, P</td>
<td>Independence from all levels essential. Some international role for common discipline</td>
</tr>
<tr>
<td>Interstate commerce</td>
<td>Constitution, N</td>
<td>N</td>
<td>P</td>
<td>Constitutional safeguards important for factors and goods mobility</td>
</tr>
<tr>
<td>Immigration</td>
<td>U,N</td>
<td>N</td>
<td>N</td>
<td>U due to forced exit</td>
</tr>
<tr>
<td>Transfer payments</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Redistribution</td>
</tr>
<tr>
<td>Criminal and civil law</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Rule of law, a national concern</td>
</tr>
<tr>
<td>Industrial policy</td>
<td>N</td>
<td>N</td>
<td>P</td>
<td>To avoid beggar-thy neighbour policies</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>Fiscal Policy</td>
<td>N</td>
<td>N,S,L</td>
<td>N,S,L,P</td>
<td>Coordination is possible</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 and internal common market</td>
</tr>
<tr>
<td>Education, Health &amp; Social</td>
<td>N,S,L</td>
<td>S,L</td>
<td>S,L,P</td>
<td>Transfers in kind</td>
</tr>
<tr>
<td>Welfare</td>
<td></td>
<td></td>
<td></td>
<td></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>Water, sewer, refuse, fire protection</td>
<td>L</td>
<td>L</td>
<td>L,P</td>
<td>&quot; &quot;</td>
</tr>
</tbody>
</table>

Note: U is supranational responsibility, ICB is independent central bank, N is national government, S is state/provincial government, L is local Government and P is non-government sectors/civil society. Source: Shah (1994).
3.2. Expanding Traditional Theory to Make it More Operational

The traditional theory of fiscal federalism offers analytic elements that are based on economic criteria above all. Yet, decentralization and fiscal federalism practices show that analysts must not be satisfied with these criteria alone. As was mentioned in Chapter 1, the reasons for decentralizing are varied and are not all—far from it—derived from economic theory. The economic approach must be competed with other criteria—management-related and sociopolitical criteria. This is the contribution of “second generation” fiscal federalism. This approach is illustrated, albeit very schematically, in Table 5 below in the form of a matrix. This interpretation guide to decentralization relies on three conceptual pillars:

(i) First, the lines in the matrix contain the analytic elements of or criteria for (de)centralization. Here, one obviously finds the economic criteria given in Table 3, but not only them. The matrix contains a second block tied to the local government managerial capacity required by decentralization, especially in developing countries and countries starting the process of decentralization. There is also a third block that aims to reveal the sociopolitical, and even demographic or historical, characteristics of the specific national context in which the (de)centralization is happening.

(ii) Second, the columns take into account the possible ways of producing public goods and services in the country studied—from the smallest institutional territory, the political commune, to the second level of government, the region, passing through the possible functional territories. The goal here is to consider, for the function in question and according to each criteria already listed, the advantages and disadvantages of each level of government. One can clearly see, from the beginning, that it is unlikely that the analysis will give a preponderant advantage to each criteria in the same column: there will be advantages to centralization (such as economies of scale) or to decentralization (such as heterogeneous preferences across the nation that are more homogenous in smaller sections of the country). They will need to be weighed against each other.

(iii) Finally, as the title mentions, the interpretation guide is repetitive. It is repeated for each function analyzed capable of being (de)centralized. The general idea is to place oneself on a continuum, starting at commune level for example and moving toward the regional level. Doing so shows that the analytic logic does not lead to one optimum dimension but to different dimensions for each function. Thus, for example, responsibility for primary education can only be analyzed by defining a production function whose inputs are teaching staff, programs, buildings, etc. Obviously, nothing says that responsibility for these inputs must be entrusted to the same level of government, or that the level in question will be the same in every country. The elaboration of school programs is entirely characteristic: in a unitary and centralizing government, programs are generally designed on national level, whereas in heterogeneous societies, local governments like to maintain a
degree of latitude to adapt the programs to local conditions by, for instance, introducing regional languages.

### Table 5 Socioeconomic Matrix of (De)Centralization by Function

<table>
<thead>
<tr>
<th>“Local – Central” Continuum</th>
<th>Commune</th>
<th>Inter communal cooperation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analytic Elements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Criteria</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Coordination</td>
<td>Private Contract</td>
</tr>
<tr>
<td>Economic:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Economy of Scale</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spillover</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Congestion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Administrative, Managerial:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Offices</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Information</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sociopolitical:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preferences</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Qualitative</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quantitative</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Legislative</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Administrative</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Efficiency of Democratic Procedures</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct Election</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voice</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

We have seen that, in decentralized systems, the governments at lower levels of the hierarchy are better able to adapt the supply of LPGs to citizens’ demands because they are better informed as to individual preferences. This information advantage, however, does not necessarily lead to an efficiency advantage. To meet their citizens’/taxpayers’ demands, local authorities must also have the ability to organize LPG production. In many cases (especially in developing and transition countries), the national government refuses to transfer powers and financial resources to lower levels of government on the pretext that their organizational capacities are insufficient without making an effort measure these capacities. But, even when the national government does everything possible to assess local capacities, such assessments can yield false results. If local capacity is underestimated, this can block the decentralization process or even strengthen centralizing tendencies. For its part,
overestimating capacity does not have fewer negative consequences: it makes the central government think that the local authorities have enough resources to take on functions beyond their real abilities, which can lead to responsibilities being transferred without material, organizational or financial support from the center. Let us also note that a lack of organizational or managerial capacity at the local or decentralized level must not be a motive to reject decentralization but, at most, a motive to delay in order to provide adequate local management training.

The decentralization matrix by function, as suggested here is a manifestation of a pragmatic approach issued from the second generation of fiscal federalism theories. It proposes an approach that brings together all those involved in decentralization. It provides a process, not results. To be coherent, this approach includes: participation by all stakeholders, elimination of strategic behaviors, transparent information on the aims and tools, coherent and accepted vocabulary, explicit weighing of advantages and disadvantages. The list of criteria is not exhaustive: what is important is that the criteria of reference be announced before analysis and explicitly defined so that they are understood by all the partners in the analysis and negotiations. Similarly, while the forms of collaboration in the “local – central” continuum are adapted to the legislation of the country analyzed, this list is not a closed list. One will, however, take care not to add too many criteria for reasons of information cost (see C5 above) and democratic transparency.

The decentralization matrix will be reproduced for each devolved function. The final objective of the process is to implement decentralization in accordance to a roadmap to be agreed with timelines, benchmarks and compliance tools (within the three F for decentralization: Functions, Finances, Functionaries).
Box 4  Subsidiarity and Decentralization

While subsidiarity is often linked to decentralization, the two terms are not synonymous. Indeed, decentralization as such does not require the government to be organized according to the subsidiarity principle. Inversely, subsidiarity can inspire decentralization programs (Fraschini, 2001). Taking a closer look, the strongest theoretical link can only be established between the subsidiarity principle and “bottom-up” decentralization. The notion of subsidiarity can be illustrated by the following schema:

Schema 3  The Subsidiarity Principle

In every more or less decentralized system of government, the current assignment of functions to the levels of authority is the result of these two antagonistic forces. Centralizing forces are usually linked to technical and rational notions (economies of scale, capacity, spillover and congestion), whereas decentralizing forces are more marked by social or democratic notions (direct participation, citizen control, appropriateness of supply to demand, etc.). In the left-hand box in Schema 3, the notion of equity between individuals is an exception to this rule: even though it tends to belong to the social arena, it reinforces centralizing tendencies. This is explained by the mechanisms of majority vote, the consequence of which is that the wishes and rights of minorities (the poor, invalids, children, linguistic or religious minorities, etc.) may not be respected at the local level, which requires intervention from the central government, often in the form of minimal standards on the quantity and/or quality of local public services.

Another problem can be seen on the right hand side of the schema and relates to inclusion of the notion of equity—this time between local government units—in the subsidiarity principle: does redistribution play a role in decentralization? The policies destined to attenuate disparities in financial capacity, needs or LPG cost between local government units and lower levels will indirectly reinforce decentralizing tendencies. Redistribution ensures local authorities have some level of protection (at least financially) against the undesirable effects of complete (or even excessive) devolution of powers. But how much redistribution is needed to respect equity and correct disparities generated by decentralization (by what
degree of decentralization?) and to ensure that the subsidiarity principle is not violated either? All in all, with the exception of functions pertaining to pure public or national goods, economic theory provides very few formulas for or advice on how to assign powers to multiple levels of government.

At the local level, the user pays principle suggests that each public service must be provided by the level of government closest to the circle of recipients. Local services such as clean water distribution, sewage and wastewater treatment, solid waste removal, and cemetery or school maintenance are typically functions that benefit the residents of a commune. Fire prevention is characterized by the distance between where firefighting equipment is stored and possible fire sites. It is therefore better to assign these to the local level: decentralization is easy to accomplish and is even desirable in terms of administrative and productive efficiency. It is also clear that pure public goods of national scope (e.g. defense, foreign relations) must be assigned to the central government and its agents. However, most government-provided goods and services do not fit easily into one single category. For these combination goods, a certain amount of decentralization (that reflects, in part, the heterogeneity of individual preferences) accompanied by coordination from the center (minimal standards to enforce equity) seems desirable. For these intermediary cases, when conflicting objectives can emerge, deciding which level of government should take responsibility for a public service can be a real brain-teaser. The debate and equally the solution are heavily guided by normative judgments and political reasoning. In conclusion, we can see that there is not one single model of how to assign public functions to the various levels of government. In practice, a wide variety of solutions can be seen around the world, each of which reflects societal preferences.\footnote{See, for ex., Ahmad, Hewitt and Ruggiero, 1997; Pola, 1999.} Overlapping responsibilities for formulating, administering and financing public policies is also common to almost all decentralized and federal systems.\footnote{In Switzerland, the federal law on the assignment of functions and the new redistribution provides outright for not only the reform of financial redistribution between the cantons but also for the untangling of public functions in nine fields that are currently shared by the cantons and the Confederation (Dafflon, 2005).} The individual manifestations of the subsidiarity principle are ultimately imbued with value judgments and the society’s social and historical trajectories.
What Type of Taxation at What Level of Government?

Governments have recourse to a very diverse range of taxes and levies to finance their public spending. When there are several levels of government, one is entitled to ask the following questions:

(i) Does fiscal federalism theory propose rules on assigning this or that tax to a specific level of government?
(ii) What happens when it is not possible to specialize taxes at different levels of government (in reality, exclusive taxes are rare) and there are therefore vertical externalities? This is the case, for example, when two levels of government tax the same taxable base.
(iii) What does a “good” tax look like at local level?

Readers will have noticed that these are not fundamental questions because underlying all of them is the question of what degree of fiscal autonomy is given to the various levels of government. One will note in passing that having an exclusive tax at local level does not in any way mean full autonomy for the local government unit receiving this tax. This question will be addressed specifically in Chapter 5.

- The traditional theory of fiscal federalism ties the issue of allocating taxes between levels of government to the functions of stabilization, redistribution and resource allocation (Boadway et al., 2000). Musgrave (1983 and 2000) defines five general principles by which to allocate taxes to this or that level of government, given that it would be an illusion to think that there could be one ideal distribution of tax revenues valid at all times and in all countries.

(P1) Taxes that contribute to macroeconomic stability must be collected by the central government. The tax revenues of local levels of government (especially communes) must be stable. Indeed, if borrowing capacity is limited or restricted to financing investments, local governments will have difficulty raising short-term loans to smooth cash flow. Another way of doing things, one that is politically more tricky, would be to create a rainy day fund when tax revenues are good that could be used during economic downturn or when tax revenues fall. With this second solution, the automatic stabilizing effect produced locally does not contradict the central government's macroeconomic actions.
The definitions presented below do not preclude the existence of other definitions and must therefore be seen more as conventions allowing one to know what one is discussing (interested readers will, for example, find that Ebel and Yilmaz’s (2001) definition of shared tax corresponds to Dafflon’s (2005: 43) notion of revenue sharing.

- **Exclusive Tax:** a tax for which only one level of government can exploit the tax base and collect all of the revenue from the tax.

- **Shared Tax:** several levels of government have access to the same tax base. If each government has full tax sovereignty, that means that each can define the tax base, and there will be as may definitions as government units, with obvious coordination and harmonization problems. A government has partial tax sovereignty when it can set part of the taxation criteria (e.g. the tax base, exemptions, deductions, tax expenditures and tax breaks, tax rate schedule, or so forth). Taxes for which the base is the same but shared by several levels of government, each of which freely sets that tax rate, enter into this category (shared tax base). Tax flexibility means that a government only sets the tax coefficient (the base, deductions and schedules are identical). Also in this category are systems that apply additional levies (piggyback taxes). Finally, mandatory taxation means that a government does not choose and must levy a tax in compliance with the rules set by a higher level of government.

- **Revenue Sharing:** Here, the tax base and tax rate are set by the highest level, which collects that tax, but a set share of tax revenues is allocated to the government units at lower levels of government. Two methods of revenue sharing exist, depending on whether sharing is based on (1) the amount of tax revenue collected from the local government in question according to the origin criterion (also called the origin principle); or (2) a distribution key that includes various elements such as the population or aims to reduce potential fiscal differences (in the last case, we speak of revenue equalization).

---

(P2) Progressive taxes that target interpersonal redistribution must remain the responsibility of the central government:

- when tax bases are potentially mobile (for example, in regard to income tax, the individuals whose resources come primarily from property and investments, from managing fortunes, rather than from salaries), or
- when one seeks to tax taxable income whose generation originates in several local government units.

(P3) Lower levels of government should tax immobile (not easily mobile) tax bases to avoid tax competition (“race to the bottom”) or, more generally, budget competition. Let us recall that taxes have a counterpart —the public services provided— and that taxpayers, according to Tiebout’s (1956) “vote with one’s feet” logic, take both aspects into account when deciding where to live.

In reality, this question is still wide open because, schematically, there are two opposing positions on the question. For some (see Wilson, 1999, for a review of writings on the subject), tax competition leads to (i) a “race to the bottom” in local taxation rates; (ii) a sub-optimal supply of local public goods; and (iii) a shift in the tax burden to the least mobile economic agents. For others, individuals’ mobility is a mechanism to reveal preferences as Tiebout claims, the efficiency of which depends...
on a relatively strict set of hypotheses. For the proponents of the Leviathan theory, tax competition is a way to limit the predatory behaviors of elected officials assumed to be opportunistic (see Brennan and Buchanan, 1977 and 1980, as the initiators and pillars of this school of thought).

**P4** Tax bases that are very unevenly distributed throughout the national territory must be “centralized”. To simplify, one can see these bases as having two origins: either endowment of local natural resources (mining, oil, etc.) or a geographical position conducive to economic development. In the first case, centralization can be justified for reasons of equity. However, this decision could create a feeling of expropriation in the regions that have these resources. But the central government can also take offense at this rent that might either be captured by a local elite or serve the interests of an opposition to the center.

In the second case, imagine that economic activities are very unequally spread across the national territory. This is obviously due to more or less favorable geographical positions in the national space, but also to local actors’ drive for development. Must one, then, follow the same centralizing logic and conclude that taxing economic activities must be the exclusive responsibility of the national authorities? Here, the situation is undoubtedly different because there is a real tension between the desire for territorial equity and the need to maintain local development incentives. More flexible and negotiated solutions (of the type redistribution or cooperation between communes) are needed in this case.

**P5** User charges and other fees levied according to the benefits-received principle (or, in practice, the "user-pays" or “polluter-pays” principle) must be used appropriately at all levels of government. This decision has a redistributive impact because the fees make it possible to have the users (and not the taxpayers) cover all or part of the cost of the service they receive. Although it raises the issue of ability to pay and the underlying redistribution policy, this type of levy is rarely envisaged and unfortunately too infrequently used in developing and transition countries.

To these five general principles, we would like to add other considerations gleaned from practical experience in tax decentralization.

1/ Local authorities should concentrate on a limited number of taxes and levies that have potentially high yields and abandon other, lesser sources of revenue. For us, obviously, it is not a matter of sacrificing to the myth of the single tax that is not desirable, if only because of the need to diversify fiscal risk! One should, however, take care to ensure that these taxes are not concentrated on an overly small number of taxpayers who, in practice, would become unavoidable actors in local budget decisions.

2/ We shall add that the be visible to the taxpayer-voters. The visibility of the tax (as a pseudo-fiscal price) requires, in exchange, the quality of public services to be visible to citizens as well or it may provoke refusal to pay the tax.
3/ Tax should not be exportable to residents outside the government unit collecting it. Tax exporting leads to shifting part of the tax burden to outside taxpayers and, thereby, weakening the link between the service provided and its financing, conducing to overproduction of the local public services.

4/ Another point that can be problematic deals with the risk of double taxation on the “horizontal” level when two fiscally sovereign government units are not able to agree on how to share tax bases. Vertical and horizontal tax coordination is not only a technical fiscal problem but, if not appropriately resolved, also raises a real problem of equity and disincentive.

5/ Local taxation must take place against the backdrop of an obligation to balance the budget, at least for day-to-day spending, which prevents the tax burden from being shifted to future generations. Visibility must be instantaneous.

4.1 Vertical Externalities and Coordination Difficulties Between Levels of Government

Local governments’ tax revenues rely not only on exclusive taxes but also on shared taxes, a source of vertical externalities. The presence of externalities is likely to influence local fiscal autonomy. A vertical (fiscal) externality is present when a fiscal decision at one level of government has an influence on another level of government’s budget constraints (Vigneault and Boadway, 1996). This is the case when:

(i) the taxes collected by one level of government give one the right to a tax credit or deduction from another level of government or;\(^{12}\) or
(ii) two or more levels of government grant tax exemptions; or even
(iii) several levels of government tax the same taxable base (called “concurrent taxation” or “tax base sharing” in North American writings on the subject).

While the first two points are largely informed in the existing literature on public finance (see, for example, Gilbert, 1996), it is only later that the theory of fiscal federalism has taken an interest in the vertical externalities that result from several levels of government taxing the same base. This may seem paradoxical since overlapping taxation (in other words, piling up tax rates on the same taxable base) concerns most countries, whether centralized or decentralized. Among other things, the possibility of utilizing overlapping taxes had already clearly been set forth more than two centuries ago by the authors of The Federalist Papers, who saw this as a way for the federal government to gain fiscal autonomy in relation to the state governments: “It is indeed possible that a tax might be laid on a particular article by a State which might

---

\(^{12}\) For example, in France, the ‘taxe professional’ (a local business tax) could be deducted from the corporate tax (a national tax). Part of the fiscal expenses paid by local taxpayers is therefore transferred in fact to national taxpayers. In Switzerland, the corporate tax on profit paid at the communal and cantonal levels is considered as ordinary business expenses and are deductible for the purpose of the federal tax on corporate profit. Thus, the vertical tax system provides a mild compensation for local and cantonal tax differentials, besides the more explicit revenue equalization policy.
render it inexpedient that [...] a further tax [...] be laid on the same article by the union,” (Hamilton, Madison and Jay, 1982).

Musgrave (1983) also addresses this issue but from a mainly administrative standpoint insomuch as he examines the appropriateness of entrusting a single collection agency with collecting all taxes levied on the same taxable base. The decision, in this case, is the result of arbitration between the economies of scale that such a solution would allow and the risk of moral hazard that it would create if the agency in question favored one level of government to the detriment of another—which is far from a hypothetical situation in countries such as Russia where local civil servants also collect the share of taxes going to the federal level (Tanzi, 1995). It was finally the articles by Flowers (1988) and Johnson (1988) that put the issue of overlapping taxation back on the agenda. Shared taxation of a taxable base by several levels of government leads to the shared tax base being overexploited (in other words, to an excessive rate of taxation).

Box 6 Overlapping Taxation and the Tragedy of the Commons: A Few Theoretical Considerations for Better Understanding

The writings of Keen (1995), Wrede (1996), Keen (1998), Flochel and Madiès (2002), and Keen and Kotsogiannis (2002) are representative of the “overlapping taxation” new approach. On the theoretical level, a certain number of conclusions, sometimes incomplete, emerge from these writings.

(i) In the simple case of two governments at different administrative levels that aim to maximize their fiscal surpluses by taxing the same taxable base, the overall taxation of the (supposedly mobile) base exceeds at equilibrium the tax rate that a single level of government would levy on the same tax base. In these conditions, total tax revenue is less that the optimum; the “vertical fiscal cooperative (Nash) equilibrium” is located in the falling segment of the Laffer curve (Flowers, 1988).

Whether the governments are benevolent or Leviathans does not qualitatively change the outcome because the vertical fiscal externalities at work are the same. If neither of the two levels of government take into account the loss in tax revenues that an increase in their own tax rate would cause to the other, then both underestimate the marginal social cost of taxation compared to the situation when the taxable base is taxed by one government only (Keen, 1995). Alternatively, if one assumes that one level of government (in practice, the central or federal level) finds itself in the position of (Stackelberg) leader, then the total taxation rate supported by the capital is higher at equilibrium than the rate supported when the two levels of government are in Nash equilibrium, but the (supposedly mobile) tax base is obviously smaller and total tax revenue is lower. The tax rate set by the central government when it is a (Stackelberg) leader is higher than the rate it would have chosen at the cooperative (Cournot-Nash) equilibrium and its tax revenues are also higher. This is obviously accomplished at the detriment of the local level that finds itself in a worse situation compared to the Cournot-Nash equilibrium. Redistributing tax revenues from the local level to the federal or national level comes with an increase in tax-related distortions (Flowers, 1988).

(ii) Interdependent relationships between local governments are not only horizontal but also vertical (Wrede, 1996; Keen, 1998; Keen and Kotsogiannis, 2002). The debate between tax competition and cooperation depends, thereby, on a comparison of the respective effects of horizontal and vertical externalities on the marginal cost of public funds (MCPF). The first, which can be shown to be all the larger and more subject to distortion when the number of competing local governments is high (Wildasin, 1988), push toward a drop in local tax rates by increasing the MCPF. The second push, rather, toward an rise in the overall tax rate on the tax base because overlapping taxes leads to an increase in the tax rate of one administrative level cutting the tax revenues of another level. Theoretical writings on vertical tax
externalities allow a certain number of results to be obtained by utilizing the theoretical framework of Zodrow and Mieszkowsky (1986) or Wildasin (1988).

These authors’ standard model must, however, be amended in two ways. Obviously one must add a higher level of government that taxes the same taxable base as the local level (the higher level can also offer the same public good to all of the population financed by taxing capital) and make the overall supply of capital endogenous (which was until now set) because otherwise the central or federal government’s fiscal policy would not be constrained and the taxes that it would raise would be equivalent to a flat-rate tax in regard to fiscal impact.13

One therefore shows (Keen and Kotsogiannis, 2002; Flochel and Madiès, 2002) that an increase in tax competition at the decentralized level strictly increases the consolidated tax revenues raised in each local government, which challenges the results obtained by Brennan and Buchanan (1977, 1980) when “horizontal” competition between local government units is not taken into account. The explanation is simple: an increase in horizontal tax competition strengthens the central government’s monopoly powers. In the extreme, one could very easily imagine a situation in which the number of competing local governments is so high (and therefore horizontal competition so sharp) that local taxation would cease and the central government would behave like an unconstrained Leviathan. If the governments are benevolent, that is to say, if they maximize the usefulness of households residing in their jurisdictions (Keen and Kotsogiannis, 2002), then intensifying tax competition at the decentralized level improves households’ welfare, at least if the public goods offered by the central and local levels are perfect substitutes for each other.14 This conclusion contradicts Hoyt’s conclusions (1991) that show that increasing the number of competing local government units lessens the wellbeing of individuals by exacerbating horizontal fiscal externalities related to the mobility of the tax base (in the Hoyt model, tax consolidation makes it possible to increase wellbeing because it makes it possible to internalize some of these fiscal externalities).

4.2. And Now a Look at “Good” Local Taxes... If They Exist

The discussion will combine arguments based on economic efficiency, equity and common sense (they are not mutually exclusive). Let us begin by eliminating taxes that are poor candidates to finance local governments (Bahl, 1999). Taxes on turnover and the value added tax (VAT) are undoubtedly among these (the issue is up for debate for large-scale local governments such as regions, and one is obliged to note that recent work calls into question the following thesis). This is mainly because exports are subject in most regimes to a nil rate at export whereas imports are taxed at the point of entry into the territory. This raises two questions: Should the local government in which the exporter resides reimburse the VAT? Can the local government that is the point of entry into the territory receive the tax revenues from taxes on imports? The VAT is, what is more, not an easy-to-administer tax at the sub-central level unless one taxes a uniform taxable base across the territory and authorizes local governments to levy an additional rate beyond the rate set by the central government—conditions that make the use of this tax unrealistic in developing countries. Taxes on retail sales also raise a problem in that there are obvious administrative difficulties collecting them...

---

13 For a more detailed discussion of this issue, in relation for instance to the incidence of vertical fiscal externalities on redistribution policies within a federation, see for example Boadway, Marchand and Vignault (1998) and Madiès (2002).

14 This is a highly restrictive hypothesis of the model and one that is never met in practice: the very sense of decentralization is that central and local public services are not the same.
from small sellers (in some countries, such taxes exist but apply almost exclusively to luxury goods).

The **business tax** is undoubtedly not a good candidate either. The major disadvantage of this tax base resides in its instability: profits are a variable that fluctuates greatly, incapable of attaining the objective of revenue stability indispensable for financing local governments—excepting the discussion above on rainy day funds. This tax is, among other things, easily manipulated via transfer prices between affiliated companies (this is especially true in developed countries), which raises the issue of the distribution of tax bases among local governments because, by definition, a tax on profits is difficult to territorialize. Obviously, it is always possible to justify taxing companies locally with an argument analogous to the benefits-received principle: the local taxes owed by businesses are a counterpart to the bundle of services that local governments provide and compensation for the costs that they must cover to bring economic activities to their territory (infrastructures, cost of urbanization, etc.).

This argument also justifies **local taxation of households**, with the difference that businesses can pass on the taxes they pay through prices upstream to production factors and downstream to consumers. Since production factor suppliers and consumers do not necessarily reside in the same local government as the businesses, taxes do not remain local, thereby generating tax exports and imports between local government units.

This possibility of diffusing taxes within the national geographic space, very difficult to trace, is an argument that goes against local taxation of businesses and justifies, for example, keeping taxation of businesses national and redistributing it to local governments according to shared distribution criteria. Among other things, even admitting that issues of incidence are resolved, the question of how to measure the advantages businesses receive from consuming the public goods offered by the local government (or, which amounts to the same thing, the costs generated for the local government by the installation of businesses) remains.

More generally speaking, the question is one of efficient sharing of local taxes between companies and households. From the standpoint of resource allocation, both income taxes on individuals and the corporation tax equally can be seen as payment for — the “fiscal cost” of — bundles of public goods and services that local governments provide. However, income taxes also include a redistribution dimension—payment according to the ability-to-pay principle — that, for its part, can be problematic when tax bases are mobile. For these reasons — mobility, and the desire to set up an identical redistribution construct for the entire nation — the **question of taxing individuals’ incomes deserves to be examined** even if, as we have seen, the economic theory of fiscal federalism states that this tax must be attributed to the central level.
Nevertheless, a local income tax would be acceptable if local governments could only collect a supplement tax at uniform proportional rate (piggyback, centimes additionnels, Steuerfuss) compared to the income tax collected by the higher level.\textsuperscript{15}

The advantages of this solution are:
(i) its simplicity of implementation (the rate scale and the tax base are the same for the entire territory);
(ii) its ability to be adjusted to the budget needs of local governments, under an obligation to balance the operating budget; and
(iii) the fact that it does not fundamentally challenge the redistribution objectives targeted by this tax.
(iv) Another advantage for most developing countries is that income tax is to a large degree a tax on official sector salaries given the difficulty determining the incomes of individual workers and income from capital.

Property tax is probably a good candidate even if it is not immune from difficulties and criticisms. First, it can only be implemented if a certain number of prior conditions are met: (1) there must be a land registry that is acknowledged, published, accessible, and updated over the long term; (2) it must be possible to identify owners and track transactions; (3) there must be updated assessments of land values, if the basic measurement is something other than plot surface area; (4) local rates must be uniform to ensure neutrality in allocating real estate between inhabitants, productive activities and the public interest; and (5) the capacity to manage this tax must exist, specifically notification, collection, conflict resolution and the legal guarantee of the amount of the tax owed on the property taxed.

This type of tax is very widely used in developing countries by urban communes but generally the subject of limitations imposed by the central government that can, in some cases, oppose the reassessment of the rateable value that serves as the taxable base. The household capitalization tax, a property tax, is, at least in theory, the tax most likely to apply the equivalence principle if the market value of the real estate holding capitalizes the services provided by local public infrastructures positively and negatively capitalizes the resulting fiscal charge. It can also be seen as partially respecting ability-to-pay since property value and management increase the capacity to contribute. The criticisms formulated against property taxes are that they are visible: residents may be reluctant to pay them, especially when they do not think of the counterpart in terms of public services. Among other things, the property owners called on to pay these taxes are often the ones who dictate the local fiscal agenda.

However, in the case of businesses, a property-based tax base can create distortions. Land or property capital is only one of a company’s production factors, and taxing only it can push the company to modify its production techniques. Choosing one factor

\textsuperscript{15} In the simplified form of taxation \((T = t \times [B - D] \times K)\) for two levels of government, central (C) and local (L), where \(T\) is the tax paid, \(t\) the tax rate schedule, \(B\) the tax base, \(D\) the tax deduction and \(K\) the coefficient for balancing the budget, we would have for the local level \(T = t^L \times [B^L - D^L] \times K^L\) that is a situation in which \(t, B\) and \(D\) are fixed by the centre; the local government decides only \(K\). This corresponds to tax flexibility in Box 5. It is the tax position of communes in the Swiss cantons.
rather than another runs the risk of setting a spatial and sectoral distribution of the tax that is not economically efficient. In addition, with the aim of equivalency, local governments cover costs related to the installation of businesses that depend on all production factors used in the production process and not immobile assets alone.

Box 7  Local Taxation in Sub-Saharan Africa

It is generally thought that local public services must, as much as possible, be financed by tax revenues raised in the territory considered. This is very rarely the case, especially in developing countries. In this regard, the financing of local governments in South Africa and Namibia (to cite two African counties) is “exemplary”, and gives a large role to local taxation (Fjeldstad and Therkildsen, 2002). The financing of South African and Namibian local governments relies mainly on own funds (more than 90% in South Africa). These funds come primarily from fees paid in exchange for services provided (electricity, water, and sanitation) but also, in urban areas, from property taxes. There are wide differences between local governments when it comes to citizens’ access to certain local public services. However, the revenue generated by local governments from fees collected on water and electricity consumption (in reality, an additional payment over the cost of providing these services) serves in practice to finance general spending. This practice amounts to endangering the close tie that should exist between fee payment and service received. For the share that exceeds cost, this practice corresponds to a hidden tax on consumption. It distorts individual choices because it calls into question the tie between payment and service received.

In addition, it would seem that in a very large number of countries (especially in Africa), local governments “tax whatever they can tax” without worrying about the economic distortions that these taxes may generate. Fjeldstad and Rakner (2003) show that the local authorities in Sub-Saharan African countries collect the fiscal product of several dozen taxes, levies of all sorts, and fees or similar (Tanzania is, in this regard, eloquent). Case studies in these countries show, among other things, that very large differences in tax rates can exist between neighboring local government units (municipalities), which leads to a large amount of contraband between neighboring local government units to avoid taxation, especially agricultural products. Finally, one very frequently sees a lack of coordination between the local authorities and the central government leading to the “duplication” of some taxes and to inconsistencies between the taxes levied by the local authorities and the export support policies implemented by the central government. One illustration is the case of Kibaha District whose taxes on cashew nut production amounted in 1997 to up to 20% of the production cost even though cashews were an export product and a very important source of currency for the country.

Using user charges and fees for financing the public sector is a continuation of the balanced budget rule, although it takes an interest in individual functions in the public sector and not the budget as a whole. This principle aims to reinstate market-type mechanisms in public sector management by introducing a direct relationship between the circle of beneficiaries and the circle of payers. The primary advantage of this approach is that it prevents excessive demand for public services and strategic “free rider” behaviors. First, we know that one budget management problem is linked to the production of services that benefit specific groups. When financed by taxes paid by all taxpayers, demand is too abundant because the causal link is broken: the beneficiaries of a service paid for by taxes do not immediately see the cost. They act as if the service were free and demand too much of it. By applying the benefit-received principle, one can avoid this: the person who receives a service (or creates the need for it) knows that he or she must simultaneously pay its cost, shared according to the advantages...
obtained. The principle includes the cost component in the demand decisions of economic agents at the very moment they express the demand. A second problem frequent in public economy is the free rider attitude. This behavior takes place when an economic agent knows he or she can receive a public service even if he or she does not reveal his or her demand for the service and without having to pay directly and individually the cost, which is distributed by taxes instead of by price. The benefits-received principle, which links price to service, prevents this strategy that would otherwise have negative allocative and redistribution consequences. With the user charges as the “price” of the public service, demand is expressed up to the usefulness to the consumer-payer: the fact that a service is not free favors a parsimonious use of resources (optimal allocation). Among other things, as the public service is paid for by users and not taxpayers, this market-type mechanism is also a criterion in redistributive equity.

The principle, however, does not apply blindly to all functions: four technical conditions must be met (Dafflon, 1998: 113 ff.).

1/ *The ability to identify beneficiaries.* For the advantages conferred by a public service to be reserved for users only if they pay the price, they must be able to be identified and individualized. Consumption is said to be divisible.

2/ *Ability to exclude.* Users that do not pay the price of the service are excluded from consuming. The ability to exclude obliges individuals to modulate the quantity they buy in function of the price. Because of this, they are obliged to reveal their preferences. The cost of exclusion via prices is added to the cost of producing and distributing the service: it must not be too high. In developing countries, this criteria obviously raises problems of access to essential services (such as water), poverty, and potential users’ financial ability to pay for services—problems that must be resolved before introducing the benefits-received principle.

3/ *Mild supply indivisibility.* Supply indivisibility complicates setting rates for public services, especially because the equal price at marginal cost rule is not applicable. Supply indivisibility means that the production of a public service makes it possible to serve several users simultaneously: once such a unit of service is produced, this same unit can be placed, more or less fully, at the equal disposition of all individuals in the group. The best known example is a television program. What real costs are generated by the tuning in of one more television set to the broadcast program? Almost nothing. Why, in this case, ban a household that derives positive satisfaction from seeing the program? The well known optimum principle under which goods and services must be sold at their marginal cost means, in this case, a zero price. However, given that such situations do not exist (here, the meaning of the adjective “mild” is highlighted), there is a marginal cost and free provision of the service is not justified (gas, electricity, telecommunications, public transit, natural resource and environmental management, water distribution, wastewater evacuation and processing, waste sorting and incineration). If, however, one admits that the total cost of producing all connected products must be financed by users’ contributions, in the absence of a marginal cost, it is appropriate to seek another rule for imputing cost. The question of setting rates
based on average cost, rather than on the marginal cost, remains open when the percentage of fixed costs is high compared to variable costs.

4/ Mild externalities. Public services for which one envisages a tariff of charges must not be the source of overly high external benefits for a large number of people other than the user-payers. Indeed, one would in this case have joint production of at least two products: one for which there is a charge paid by consumers who express a demand for the product, and the other that has more or less pronounced characteristics of a pure public service for which no charge is set, which complicates considerably the formulation of equitable prices. Strong externalities mean that a service benefits, first, to the entire local government and, to a lesser extent, individual consumers. In this case, it is unreasonable to ask the latter to cover the total cost of the service via causal contributions. This would not meet the demands of either equity among individuals or optimal resource allocation. In the case of a positive externality, the service would be under-produced, as private decisions would in this case not take into account positive social effects. Public finance theory proposes correcting this situation through grants financed by taxes.

Like what we have seen for the assignment of powers to levels of government, and from the same author (Shah, 1994), Table 6 reproduces a possible view of the assignment of fiscal resources, taxes and user charges to levels of government. The comments, however, sometimes hint at a sharing of taxing powers between the regional and local levels without specifying the nature of this sharing (which is not self-evident and is subject to debate: see Box 5). At other times, the author assumes or mentions the need for an immobile tax base. Vertical harmonization and implementation cost arguments also emerge.
### A Representative Assignment of Taxing Powers

<table>
<thead>
<tr>
<th>Types of Tax</th>
<th>Determination of Base</th>
<th>Collection and Rate</th>
<th>Administration</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customs</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>International trade taxes.</td>
</tr>
<tr>
<td>Corporate income</td>
<td>F, U</td>
<td>F, U</td>
<td>F, U</td>
<td>Mobile factor, stabilization tool.</td>
</tr>
<tr>
<td>Resource taxes</td>
<td></td>
<td></td>
<td>F</td>
<td>Highly unequally distributed tax bases.</td>
</tr>
<tr>
<td>Resource rent (profits/income) tax</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>Benefit taxes/charges for state-local services.</td>
</tr>
<tr>
<td>Royalties, fees, charges; severance taxes; production, output, and property taxes</td>
<td>S.L</td>
<td>S.L</td>
<td>S.L</td>
<td>To preserve local environment.</td>
</tr>
<tr>
<td>Conservation charges</td>
<td>S.L</td>
<td>S.L</td>
<td>S.L</td>
<td>Redistributive, mobile factor, stabilization tool.</td>
</tr>
<tr>
<td>Personal income</td>
<td>F</td>
<td>F, S</td>
<td>F</td>
<td>Redistributive.</td>
</tr>
<tr>
<td>Wealth taxes (taxes on capital, wealth, wealth transfers, inheritances, and bequests)</td>
<td>F</td>
<td>F, S</td>
<td>F</td>
<td>Benefit charge, e.g., social security coverage.</td>
</tr>
<tr>
<td>Payroll</td>
<td>S</td>
<td>F, S</td>
<td>F</td>
<td>Border tax adjustments possible under federal assignment, potential stabilization tool.</td>
</tr>
<tr>
<td>Multi-stage sales taxes (value-added tax, [VAT])</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>Higher compliance cost.</td>
</tr>
<tr>
<td>Option A</td>
<td>S</td>
<td>S, L</td>
<td>S, L</td>
<td></td>
</tr>
<tr>
<td>Option B</td>
<td>F</td>
<td>S</td>
<td>F</td>
<td></td>
</tr>
<tr>
<td>&quot;Sin&quot; taxes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excises on alcohol and tobacco</td>
<td>F</td>
<td>S, L</td>
<td>S, L</td>
<td>Health care a shared responsibility.</td>
</tr>
<tr>
<td>Betting, gambling</td>
<td>S</td>
<td>S, L</td>
<td>S, L</td>
<td>State and local responsibility.</td>
</tr>
<tr>
<td>Lotteries</td>
<td>S</td>
<td>S, L</td>
<td>S, L</td>
<td>State and local responsibility.</td>
</tr>
<tr>
<td>Race tracks</td>
<td>S</td>
<td>S, L</td>
<td>S, L</td>
<td>State and local responsibility.</td>
</tr>
<tr>
<td>Taxation of &quot;Bads&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carbon</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>To combat global/national pollution.</td>
</tr>
<tr>
<td>BTU taxes</td>
<td>F, S, L</td>
<td>F, S, L</td>
<td>F, S, L</td>
<td>Pollution impact may be national, regional, or local.</td>
</tr>
<tr>
<td>Efficient charges</td>
<td>F, S, L</td>
<td>F, S, L</td>
<td>F, S, L</td>
<td>To deal with interstate, intermunicipal or local pollution issues.</td>
</tr>
<tr>
<td>Parking fees</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>To control local congestion.</td>
</tr>
<tr>
<td>Motor vehicles</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Registration, transfer taxes, and annual fees</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>State responsibility.</td>
</tr>
<tr>
<td>Driver's licenses and fees</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>State responsibility.</td>
</tr>
<tr>
<td>Business taxes</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>Benefit tax.</td>
</tr>
<tr>
<td>Property</td>
<td>S</td>
<td>L</td>
<td>L</td>
<td>Completely immobile factor, benefit tax.</td>
</tr>
<tr>
<td>Land</td>
<td>S</td>
<td>L</td>
<td>L</td>
<td>Completely immobile factor, benefit tax.</td>
</tr>
<tr>
<td>User charges</td>
<td>F, S, L</td>
<td>F, S, L</td>
<td>F, S, L</td>
<td>Payment for services received.</td>
</tr>
</tbody>
</table>

**Note:** U is supranational agency, F is federal, S is state or province, L is municipal or local.

**Source:** Shah (1994).
Case studies in Sub-Saharan African countries (in particular, Tanzania and Uganda) show that there is often resistance to taxes at the local level, or even fiscal revolt as was the case in 1998 in Arumero District in northeastern Tanzania (Kelsall, 2000), and that the local authorities resorted to brutal measures to collect taxes (use of local militias or “commandos” entrusted with tax collection), to such an extent that, in 2002, the central government was obliged to issue a directive ordering the local authorities to refrain from such behaviors. This resistance to taxation is particularly marked in these two countries for the capitation taxes that are owed by men (women are exempt) and therefore penalize the poorest. Fjeldstad (2002b) notes, among other things, that if coercion makes it possible to increase the amount of tax per capita (at least in the short term), in reality it only materializes (and strengthens) citizens’ lack of trust in local authorities and the poor quality of the services offered in exchange.

In this regard, we shall emphasize that when advocating more fiscal autonomy in the name of increasing the accountability of local authorities (efficiency argument) one must be careful not to intensify coercive behaviors. The question of tax collection also concerns the taxpayers themselves. Trust requires three levels:

1. trust in local governments that use tax revenues to provide the expected public services;
2. trust in local governments to take “honest” measures likely to collect taxes and provide public services “honestly”; and
3. trust in other citizens who must share the tax burden fairly (with the example role of representatives).

Fjeldstad and Rakner (2003) report experiments in some African countries (Tanzania and Uganda during the 1990s) aiming to grant the tax administration a certain degree of autonomy (the idea was to reduce daily interference with the political authorities as much as possible). It clearly emerges that although the amount of taxes collected (compared to GDP) increased initially, it was temporary (this phenomenon was also seen in several Latin American countries). There are two reasons for this. First, the tax administration became a clearly identified target in light of the remuneration offered and possibility of seizing rents. Very interesting (in Uganda) is the fact that the tax collection authority’s reputation worsened once it was no longer able to attain its tax revenue objectives. (In this regard, we shall insist on the potentially perverse role of donors and the Ministry of Finance that, by raising tax collection objectives, contributed in reality to weakening the credibility of the agency that could not meet these objectives.) In some cases, such as in Mozambique, the first stirrings of privatization have been seen for customs collections.
What Autonomy?

Local governments’ autonomy—which is the very expression of their ability to decide and act in terms of the devolution of powers—raises serious conceptual problems. Definitions are diverse and often refer to a specific aspect of autonomy. Autonomy can only be relative. Efforts at clarification are therefore needed. Autonomy has two components: financial autonomy that concerns local governments’ resources, and budget autonomy that concerns their spending. The link between financial and budget autonomy is not lacking in ambiguity. On one hand, financial autonomy conditions budget autonomy, in virtue of the budget accountability principle. This principle says that local governments must shoulder the expensive (fiscal) consequences of their spending decisions or, alternatively, that they can finance functions only if they have the resources to do so. On the other hand, it is clear that the nature of the functions fulfilled by local governments also determines the nature of the revenue that finances them, the logical sequence being first define the powers devolved to local governments, then—and also—give them the means to shoulder them. This logic can be illustrated by Schema 4 showing the decentralized budget framework in the two upper boxes highlighting the assignment of powers and resources.

The sought-after definition of autonomy (financial, budget) and its corollary in terms of budget accountability inevitably leads to the question of the imbalance between the two sides of the scale when gaps form between revenue and spending or when the institutional assignment of responsibilities and resources does not coincide. This requires one to address the issue of transfers, which often account for more than half of local government units’ revenue in developing countries and that are often accused of limiting the autonomy of recipient local governments. The question is, unfortunately, not that simple. What is more, one must be able to tell these transfers from other types of resources such as those from sharing (the product) of taxes (such as revenue sharing).

Schema 4 compares, in a decentralized operating budget, the spending engendered by the attribution of responsibilities (on the left) and possible revenue. In the upper right quadrant, taxation itself is in function to own spending. In other words, the criterion of budget accountability says that local government units that determine their own functions finance their spending decisions with their own taxes. However, the functions delegated in a “principal – agent” relationship would be financed by financial transfers (grants and subsidies) according to the adage “he who orders pays.”

---

16 Here, we shall only mention the best known political definition of autonomy, widely accepted by developed countries and transition economies: The European Charter of Local Self-Government and its explanatory report, 1986, reprinted 1996, Council of Europe Publishing, Strasbourg.
second part of the schema identifies, in the lower left box, the possible causes of gaps. The solutions are found either in reflections leading to a new assignment of powers (number 2 in the left margin) or in financial transfers the characteristics of which must be clarified (bottom right box).

Schema 4 The Decentralized Budget: Powers and Gaps

<table>
<thead>
<tr>
<th>Decentralized Budget</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spending</td>
</tr>
<tr>
<td>Revenue</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Distribution of Functions and Powers</th>
<th>Distribution of Resources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Own choice of LPGs</td>
<td>Own Taxation</td>
</tr>
<tr>
<td>Specific Functions (devolved or delegated)</td>
<td>- shared tax(es)</td>
</tr>
<tr>
<td>Delegated Functions (agency) (1)</td>
<td>- exclusive tax(es)</td>
</tr>
<tr>
<td>- in general</td>
<td>- choice between taxes and user charges</td>
</tr>
<tr>
<td>- incentives</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Supply Gaps</th>
<th>User Charges</th>
</tr>
</thead>
<tbody>
<tr>
<td>- gap between functions and resources</td>
<td>- grants not attributed to tax sharing</td>
</tr>
<tr>
<td>- tax system uniformity</td>
<td>- specific incentive grants</td>
</tr>
<tr>
<td>- budget deficit</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Production Gaps</th>
<th>Gap in Financial Resources</th>
</tr>
</thead>
<tbody>
<tr>
<td>- spillover and congestion effects</td>
<td>- additional budget grants</td>
</tr>
<tr>
<td>- disparities in cost</td>
<td>- idem.</td>
</tr>
<tr>
<td>- idem., or a coefficient system</td>
<td>- corrective grants</td>
</tr>
<tr>
<td></td>
<td>- cost redistribution</td>
</tr>
</tbody>
</table>

Notes:

(1) The associated spending can take three forms:
- vertical delegation (functions delegated by the center to intermediary governments, and by the latter to local governments (merit LPGs, standards);
- horizontal delegation (functions attributed by one commune to another by contract or as part of a cooperation agreement of the single or multi-purpose inter-communal syndicate type);
- day-to-day spending (interest and amortization of the infrastructure tied to the introduction of standards).

(2) Here, the solution to favor is to re-examine the assignment of powers and/or resources, grants being only subsidiary ways to bridge the remaining gaps.
5.1. Financial Autonomy

We speak of financial autonomy “when a local government is able to obtain through its own means the financial resources it needs without recourse to or depending on other local governments situated at a higher level or the same level of government” (Dafflon and Perritaz, 2003).\(^{17}\) Independent access to financial resources (decentralized taxes, user charges, revenue from property) is, for local governments, the crux of the matter: only this autonomy can guarantee their ability to make choices in response to local preferences. Here too, financial autonomy should not be absolute. Rather, it is the volume of resources that a local government can devote to financing its own decisions that is important.

Financial autonomy must not be confused with fiscal sovereignty (Box 5). The second corresponds to “the right to invent taxes, that is to set tax bases, the circle of taxpayers, the method of calculation and the rate schedule, as well as the right to levy the tax; in other words, to collect and manage, or at least be the first to do so, fiscal disagreements” (Dafflon, 2005b: 43). Yet, a local government needs only the ability to set its own tax coefficient (piggyback tax) to obtain its own resources. Financial autonomy can make due with fiscal flexibility alone. Similarly, even when a large share of resources comes from transfers from a higher level of government, it would be hasty to conclude that there is a high degree of dependency. Before reaching such a conclusion, one must analyze the transfer’s institutional nature and technical characteristics. For instance, when the Swiss cantons ceded to the Confederation sharing of direct taxation on individual incomes and company profits in 1934, they demanded in exchange that a 30% share of these tax revenues with no restrictions on use of the funds be written into the Constitution. Therefore, formally, there was indeed a vertical financial transfer that was not counted as part of the cantons’ own resources when calculating financial autonomy ratios (see below), even though from the historical and constitutional standpoint, this (new) 17% share of the direct federal tax is legally owed to the cantons without restrictions and the Confederation cannot decide otherwise—unless it modifies the Constitution.\(^{18}\)

- The financial autonomy of regional and local levels of government can be measured in different ways. The first consists of comparing local governments’ own tax revenues to their total revenues (OECD, 1999; Blöchliger and King, 2006). The OECD calls this measurement fiscal autonomy. From this standpoint, fiscal autonomy is a sub-set of financial autonomy because it consists of only the part that concerns local taxes to the exclusion of other own resources. This obviously does make sense only inasmuch as one implicitly admits that local governments have more latitude with regard to their tax revenues than for non-tax revenues. Yet, this is not necessarily the case. Along the same lines, comparing local governments’ tax revenues to the public administration’s total tax revenues (understood as the total mandatory levies collected

---

17 Guenguant and Josselin (2005: 150) speak of “capacity to raise revenue autonomously.”
18 Prior to 2008, the cantons received 30% of the FDT. Since 2008, with the introduction of the new system of fiscal equalization, the percentage has fallen to 17%. The old 13% equalization share is now part of the new revenue equalization scheme. The modification was voted in on 28 November 2004. DFT sharing with the cantons therefore lasted from 1934 to 2007 at a 30% and 17% since 2008.
by public administrations at all levels) is often proposed. This measurement runs up against the same problem as the previous one because local governments’ “room for maneuver” is in fact very often limited by higher levels of government (including for their own tax revenues). By distinguishing between fiscal autonomy and financial autonomy, these definitions create a risk of confusion with the concept of fiscal sovereignty, as we shall see below.

The OECD proposes other measurements: for instance, calculating the share of total resources occupied by transfers in sub-central levels of government (transfers over which they very often have little say, especially when they take the form of specific grants—see below). The fundamental question that is raised is, in reality, knowing over which of their own resources local governments have discretionary power or latitude. We shall insist in passing that, unfortunately, there is very little correlation between all these indicators of fiscal autonomy (which raises serious methodological problems for empirical research that involves decentralization variables).

- For the OECD (1999) and Blöchliger and King (2006), the degree of fiscal autonomy that a local government has is greater when they are free to modify their tax rates and define their tax bases without another level of government setting a ceiling on rates or the amount of tax revenues collected from a specific tax. For Dafflon (2005: 43), the concepts have become confused: fiscal autonomy and fiscal sovereignty do not designate the same thing. The OECD’s definition corresponds, in his terminology, to partial fiscal sovereignty, or even simple fiscal flexibility. That is to say, a local government’s ability to determine the contours of the taxes it can collect. But this still does not mean that exercising fiscal sovereignty gives it much fiscal or, more generally, financial autonomy. Thus, in many developing countries and transition economies, property tax is assigned to the intermediary or local levels of government, leaving them the job of defining its contours: there is a certain degree of fiscal sovereignty (flexibility). But, as the revenue from this tax is generally little compared to total resources, it would be dangerous to conclude that this same tax gives them significant financial autonomy! Inversely, we can conceive of a situation in which the central government fully determines the outlines and modalities of a local tax and local governments must collect it—they have no fiscal sovereignty. But the revenue from this tax belongs entirely to them: if it is considerable, it may even provide a certain degree of financial autonomy. Fiscal autonomy and fiscal sovereignty are not interchangeable concepts.

A tricky question deals with the fact that, in a very large number of countries, the tax revenue from taxes collected by one level of government is shared between levels of government (revenue sharing – Box 5). In this system, the tax base and taxation rates are generally (but not necessarily) set at national level (or at least by a higher level of government). The degree of autonomy depends, in this case, on how much latitude local governments have to negotiate when defining the formula that will be used to distribute tax revenues between levels of government. We note, however, that the shares received or paid by local governments (if the tax is collected by the local level that turns part of it over to the higher level) under revenue sharing, if they are general taxes and acquired by right (for example, written into the Constitution), only bring
about a relative loss of autonomy compared to the share of revenue that is allocated or to specific grants (Dafflon and Perritaz, 2003).

The OECD utilizes a classification system that distributes taxes in (decreasing) function of the degree of control that local governments have over their tax revenues. This classification makes it possible to compare OECD countries. However, some of these indicators are quite simply not available for developing countries (or are incomplete). The options available in OECD nomenclature are as follows:

(1) sub-central governments (SCGs) determine only their tax base;
(2) SCGs freely chose their tax rate;
(3) SCGs freely chose their tax rate and taxable base;
(4) tax revenues are shared, which implies:
   (a) that SCGs are free to determine revenue sharing;
   (b) that they must automatically be consulted in the case of changes to the tax revenue distribution key;
   (c) that the distribution of tax revenues is determined by a legal standard that can be challenged unilaterally by the central government or more generally by a higher level of government;
   (d) that distribution many unilaterally be changed and revoked by the central government;
(5) the central government (or a higher level of government) sets the tax rates and tax bases allocated to SCGs.

- To conclude this section on fiscal autonomy, let us mention three additional points:

1/ The higher level of government can define a system of tax deductions that is imposed on local governments, in all or part of the country, which reduces by as much the local governments’ fiscal sovereignty and autonomy (it even happens that the higher level of government—often the central government—offsets this lost revenue by covering this lost revenue and compensating, in the form of transfers, the local governments in question—which causes an additional loss of financial autonomy).

2/ An important point that is often neglected but of consequence in fiscal autonomy is obviously the more or less pro-cyclic nature of the fiscal revenues that local governments have (see section 4.2 on local taxation the bases that are more or less sensitive to the current economic situation). Similarly, it is clear that the buoyancy of tax bases must be taken into account (in other words, when the tax is more or less productive). Also along the same lines, one must not neglect the concentration of the tax base on a limited number of taxpayers because this introduced an additional risk factor for the local government in question.

3/ Tax competition between local government units, such as yardstick competition, has an impact on local governments’ fiscal autonomy. If competition causes them to lower the tax coefficient, for example, local governments use their fiscal flexibility (downward adjustment), but this reduces their fiscal (and financial) autonomy by reducing the proportion of their own resources in their total revenues (unless in the
medium term they find a larger tax base as a result of lowering the coefficient). Along the same lines, tax base sharing (Box 5) by several levels of government has an effect on local governments’ fiscal autonomy insomuch as it generates vertical externalities between levels of government (in the same way that the vertical deductibility of a local tax from the base or amount of tax to pay to a higher level—for instance the central government—shifts part of the tax burden to the national taxpayer and therefore calls into question the principle of local accountability).

5.2. Budget Autonomy

Budget autonomy concerns the capacity of a local government unit to “decide alone, fully independently, the categories, quantity and quality of services that it intends to offer its residents” (Dafflon and Perritaz, 2003). It is very frequently measured in OECD studies by the ratio of sub-central levels of government to total public spending by all public administrations (this measurement is also seen as an indicator of spending decentralization).

In reality, things are more complex because local governments’ budget autonomy cannot be summed up by this indicator alone. Two categories of problems can arise: one concerns the nature of the functions fulfilled by local governments; the other is a matter of law via the normative density of service obligations. Furthermore, these two categories partially overlap.

First, let us recall that the functions fulfilled by local administrations can be classified in three categories: (i) deconcentrated functions fulfilled as an agency, (ii) delegated functions, and (iii) devolved functions.

(i) For deconcentrated functions, local governments fulfill functions and provide services without having any (or having only very little) freedom to determine the level and quality of the public services produced. They therefore fulfill the function of agents at the service of the central administration or a higher level of government (e.g. paying certain benefits or remunerating civil servants). In a way, we can say that the central government dictates supply, and the local governments manage production according to criteria determined by the central ministries concerned.

(ii) For delegated functions, local administrations have only partial influence over the level and quality of services they provide. The central government entrusts local governments with responsibilities in fields that it wants to control in part but for which it knows that proximity of production is important because of the diversity of situations within the country. But, it frames this delegation with conditions, standards or supervisory norms to guarantee a minimum level of harmonization between regions of the country. For instance, this is the case with primary schools in a large number of developing countries where local administrations accomplish the functions that they are assigned but do so based on standards and rules, and where the salaries of teachers and numerous other lines of spending are financed and controlled by the central administration.
(iii) Devolved powers are the true arena of budget autonomy. In this case, local administrations have strong controlling power and full responsibility for services and the quality of these services.

The second category of problems comes from the normative density of central legislation. These are the legal provisions adopted by the center but that only affect the communes’ budgets. They are not deconcentrated functions: local governments fulfill them for themselves, not for the center. Communes have a great deal of freedom in fulfilling functions; production is not supervised but local governments must meet the targets set by the center. They can be seen as a form of delegated function. For instance, in Switzerland, the Confederation sets environmental policy targets, the cantons coordinate them, and the communes must meet the targets and finance them. In the OECD classification, we would have a “measure of decentralization” very heavily in favor of the communes. In reality, none exceed the threshold of the standards set by the center. They exercise no budget autonomy in this field. What is more, the user pays and polluter pays principles function fully for such functions.
6

Budget Transfers

In developing countries, budget transfers are one of the primary sources of revenue for local governments. They account for more than half of the resources of local administrations in many of these countries. We first propose a set of justifications for these budget transfers by showing that there is a difference in the approaches taken in traditional theories of fiscal federalism that emphasize mainly the need to correct externalities between levels of government (vertical externalities) and between local governments belonging to the same administrative level (horizontal externalities), and those taken in the “second generation” theories that emphasize the risk that these transfers weaken the autonomy of local elected officials (with, as a corollary, a weakening of the accountability principle) and make these elected officials less concerned with fostering growth in their territory (rent seeking, a potential source of corruption, is then fostered). We then propose a classification system for these transfers by showing that it is often difficult to distinguish between budget transfers and revenue sharing. Finally, we propose a few principles on the proper use of budget transfers.

6.1 Justifications for Budget Transfers

Economic theory has advanced several reasons to justify the existence of budget transfers between levels of government and between local governments. They are summarized in Table 7. We can also highlight them by examining Graphic 4 on decentralized budgets.

(i) Transfers aim to compensate the local level when it is obliged to fulfill an agency function for a higher level of government, such as providing a basic service throughout the territory in compliance with the demands of the central government.

(ii) Transfers also aim to correct vertical gaps. This is the case in particular when local governments’ own resources are insufficient for them to finance the spending under their responsibility.

(iii) Transfers can, finally, aim to correct externalities (such as spillover) linked to the provision of certain local services. This is the case in particular when the area of consumption of public services exceeds the borders of a given local government unit or when non-resident economic agents can benefit from the services provided by a local government without paying the cost. For example, the education services provided and financed by one given region have positive externalities on other regions (especially when individuals are mobile). The problem is that local governments have a tendency to ignore both the positive and negative externalities that they generate for
the rest of the economy when they set the level of public goods and services that they wish to offer their citizens (the problem of coordination). This can lead to a sub-optimal situation from the standpoint of the national government (under-provision of services with positive externalities and over-provision of goods and services that have negative externalities on the rest of the economy).

(iv) Finally, budget transfers also aim to correct horizontal gaps between local governments. In this case, transfers play a role in resource redistribution (and also redistribution of needs). They generally aim to “balance out” (i.e. bring closer to the average) the resources available to local governments to that they can offer comparable quality (in the same price range) public services in a given territory. Additional resources are, in this case, transferred to local governments that have capacities to contribute below the national average or higher needs than the national average (for example, communes in mountain zones).

<table>
<thead>
<tr>
<th>Table 7</th>
<th>The Economic Justification of Inter-Government Budget Transfers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Vertical Fiscal Gap</strong></td>
<td></td>
</tr>
<tr>
<td>- gap between functions and resources</td>
<td>new assignment of functions/resources, fiscal decentralization or tax sharing</td>
</tr>
<tr>
<td>- budget deficit</td>
<td>new assignment of functions (long-term solution) or tax sharing</td>
</tr>
<tr>
<td>- the center seizes the tax</td>
<td>tax sharing, increasing the local levels’ fiscal sovereignty</td>
</tr>
<tr>
<td>- excessive federal/central tax rate</td>
<td></td>
</tr>
<tr>
<td>small or weak tax base, in a uniform tax system</td>
<td>fiscal decentralization or tax sharing</td>
</tr>
<tr>
<td>limited or few tax revenues (when one accepts a certain degree of tax competition between local government units)</td>
<td>material tax harmonization (cessation of tax competition)</td>
</tr>
<tr>
<td><strong>2. Spillover Effects</strong></td>
<td></td>
</tr>
<tr>
<td>- the local authority does not take into account the benefits/costs generated for other local governments → sub-optimal/over-optimal LPG production</td>
<td>reorganization of preexisting territorial structures to establish a correspondence between the institutional and functional dimensions</td>
</tr>
<tr>
<td><strong>3. Redistributive Fairness Between Individuals</strong></td>
<td></td>
</tr>
<tr>
<td>- social insurance, health, education, social aid, allocated at the regional/local level for reasons of efficiency</td>
<td></td>
</tr>
<tr>
<td><strong>4. Minimal Standards / Merit Goods / National Priorities</strong></td>
<td></td>
</tr>
<tr>
<td>- standards dictated by the center to facilitate production factor mobility / to defend the interests of minorities</td>
<td></td>
</tr>
<tr>
<td>- imposition of the center’s preferences on the local level in the interest of national wellbeing</td>
<td></td>
</tr>
</tbody>
</table>
5. Horizontal Fiscal Gap

- inequitable distribution of natural resources
- gaps in financial capacity
- differences in costs/needs, opportunity to generate an economy of scale, for demographic, topographic, socioeconomic, etc. reasons

<table>
<thead>
<tr>
<th>New assignment of functions/resources, inter-governmental collaboration</th>
<th>general lump-sum grants (sharing, redistribution)</th>
</tr>
</thead>
<tbody>
<tr>
<td>specific matching grants, if the redistribution objective combines with the demand for merit goods (correction or compensation)</td>
<td></td>
</tr>
</tbody>
</table>

6. Macroeconomic Policies

<table>
<thead>
<tr>
<th>Stabilization: encourage local spending during recessions and discourage it during economic upturns</th>
<th>encourage private sector participation</th>
<th>Recession: specific lump-sum grants (for investments) (incentive)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regional Development: maintain or foster economic activities in the outlying regions that have development potential</td>
<td>help establishing framework conditions; regional sectoral policies to encourage the private sector; support for new activities</td>
<td>sectoral grants (e.g. mountain agriculture, tourism, LIM) (incentive and support)</td>
</tr>
</tbody>
</table>


The above justifications for budget transfers between governments very often forget a fundamental point raised, among others, by Singh and Srinivasan (2006: 5): “the standard public finance question takes subnational jurisdiction’s income as given and looks at the incentive effects of tax assignments and transfers. The second generation of fiscal federalism (SGFF) growth perspective examines the effects of the tax and transfer system on incentives to increase income (e.g., through public or private investment) [...] the allocative efficiency of the tax system in a standard public economic sense is of second importance to fiscal autonomy on the revenue side.” Weingast (2006: 15) emphasizes, along the same lines, that the major weakness of budget transfer policies (in which he includes “tax sharing”) particularly in developing countries resides in their low incentive power in regard to local growth (to which one must add the risks of corruption). One example is a stylized transfer system whose distribution formula includes the economic and demographic characteristics of potentially eligible territories (revenue and population, to simplify). We assume that this formula cannot be changed from one year to the next. In other words, only the amount of the fund devoted to transfers can change from one year to the next. We assume that the country is made up of N regions and that each of them receives 1/N of the fund (R) devoted to transfers. The average amount of the transfers received by each region therefore equals R/N. We now assume that one of the regions’ tax revenue increases by one unit (thanks to regional growth). The total amount of the transfer fund is then equal to R+1, and the average transfer amount becomes (R+1)/N. In other words, the region with the most growth only receives 1/N of the tax revenues generated in its territory while it covered the entire cost of the public policies (infrastructures, for example) that allowed it to increase its growth. Furthermore,
Caraga and Weingast (2003: 17) call the often dis-incentive nature of budget transfers the “fiscal law of $1/N$”. This is also true of transfers distributed systematically on the basis of the population.

A certain number of empirical studies have attempted to measure the “marginal retention rate” for the revenue generated in a given territory (Weingast, 2006). This retention rate was evaluated in the mid-1990s at 23.3% in Mexico (at certain periods, it was shown that all of the tax revenue generated in the Mexican states was put into a common pot so that the marginal retention rate was more or less equal to $1/33$ (for 33 states)). In India, the marginal retention rate was evaluated as being in a range of 20% to 30% (the same magnitude was obtained in Pakistan). The same type of calculation was done for Russian cities, and the result was on the order of 10%. In other words, this means that for each ruble generated locally, the government of the region in which the city is located “captures” most of this revenue increase by lowering its transfers by 90 kopeks! In contrast, a study conducted in China (1981-1992, a period of strong growth) shows that Chinese provinces retained on average 89% of the additional tax revenues generated in the provinces and that 68% of the provinces had a marginal retention rate of 100%.

One way of reconciling the objectives of both redistribution (which should not be called into question) and incentive (in terms of growth) is to define a transfer policy that is not linear. An example makes it possible to understand the underlying principle. We assume that the total tax revenues generated in a given province equal 11 billion dollars and that the central or federal government retains 75% and redistributes 25% to the province in question. The federal government receives 8.25 billion whereas the province receives 2.75 billion. The province obviously has very little incentive to “produce growth”. Now let us assume that the rule is changed as follows: the federal government retains 80% of total tax revenues collected in the province up to a ceiling of 10 billion and transfers 20% to the province. Above 10 billion, the central government transfers 75% of the collected tax revenues to the province and retains only 25%. The 11 million are distributed the same between the federal government and the province (8.25 to the first and 2.75 to the second). The essential difference between the two transfer methods is that the province’s marginal retention rate is higher in the non-linear case. Let us continue the example by assuming that growth in the province increases by 10% per year for five years. It is in this case easy to show that under the traditional transfer schema, the province’s resources will increase by 1.68 billion, compared to 5.04 billion with the non-linear method. After five years of growth at 10% per year, the province’s revenue increases by 61% in the traditional scenario and by 183% when the marginal retention rate is high (non-linear scenario). The federal government also has more resources that it can allocate to its redistributive target.
6.2 Criteria Definitions

A grant typology is tricky for two reasons. First, the vocabulary is neither clear nor unequivocal. While English writing is thought to be more precise in this area and seeks to be a reference, King (1984: 87-88) enumerates several uses of different terms for grants with the same characteristics, as well as several identical designations that have different contents. The French-language equivalents are approximations, with one French term sometimes corresponding to several English designations. All discussions related to grant systems should first resolve this aspect: without precise definitions, confusion risks winning out over coordination. Wiseman (1990) mentions numerous criteria and insists heavily on this point for a second reason. The vocabulary, full of nuances, hides very subtle variations in objectives and effects. Grant systems that address the same local governments and seem contradictory if we take them separately must be analyzed in an overall manner and are, in fact, the outcome of delicate political arbitrations. In other words, a grant formula that favors a certain group of local governments becomes honorable and acceptable because it allows these local governments to accept another, less favorable, formula, whose consequences it attenuates. Their complexity and complementary nature (positive or negative) is what makes them acceptable!

Now that this warning has been issued, we can attempt to retain and clarify the habitual grant typology. It proposes a few crucial distinctions in regard to the nature of grants. We have selected three: use, method of calculation, and fund availability. They can be combined as shown in Schema 5.

6.2.1. According to Use: General or Specific

- Grants are **general** (block, or unconditional) if the amount is **freely available** to the recipient local government. In general, tax sharing belongs to this category.

- Grants are **specific** (selective, conditional, restricted, or categorical) if selection criteria limit the use of the funds by **allocating them to a function** or to spending specified by the government that issues the grant. The allocation can be more or less precise (for example, a grant for secondary education teaching is probably less restrictive than a grant to develop the use of computers in secondary schools).

6.2.2. According to the Method of Calculation: Matching or Non-Matching

- **Matching** grants are conditional on or proportionate to a specific expenditure when the amount received varies in function of either the amount of subsidized spending or the amount of own funding that the recipient local government is ready to devote to the subsidized function.

- Grants are **non-matching** (or lump-sum) when the local government receives a set amount to spend (for the specified function if the grant is specific, but not necessarily), irrespective of its own contribution; or, in the case of tax sharing,
when the amount is pre-determined and guaranteed in proportion to tax revenues.

**Schema 5  Grant Typology**

<table>
<thead>
<tr>
<th>use</th>
</tr>
</thead>
<tbody>
<tr>
<td>method of calculation</td>
</tr>
<tr>
<td>fund availability</td>
</tr>
<tr>
<td>expected result</td>
</tr>
</tbody>
</table>

6.2.3. **According to Fund Availability: Closed-Ended or Open-Ended**

- Grants are **closed-ended** if the total budgetary means made available by the government issuing the grant for a subsidized function are **limited**. In this case, selection criteria must be determined in case total individual demands exceed the available resources.

- Grants are **open-ended** when the situation is the opposite of what is described in the preceding paragraph. In general, this means that a local government can receive a grant once certain conditions or criteria are met: it has a right to the grant independently of the issuing government’s financial situation and the latter cannot use its budgetary situation or other circumstances to refuse or reduce the grant.

To be complete, two notes on “general” and “specific” grants must still be made here:

(i) Even the attribution of a general grant can be tied to compliance with one or more “conditions”. For example, it can vary in function of an effort to make or a budget
threshold to follow (minimum fiscal effort, volume of the beneficiary’s own spending, or a set percentage of the recipient local government’s tax revenues).

(ii) All matching grants are “discriminatory” in that they favor local governments with strong financial capacities, to the detriment of less well-off local governments. While the first are able to make a considerable contribution to financing the project and, consequently, can receive large grants, the latter have difficulty meeting the minimum co-financing conditions and are therefore often excluded from the circle of potential recipients, even if their needs are greater than those of the competing local governments. In other words, the local governments that receive grants are not equal in the face of the residual spending necessary to finance the balance of the expense after the grant. To avoid the downward spiral of horizontal gaps, inequality of access must be neutralized by a system of equalization transfers.

(iii) Another unwanted effect of matching grants is that they oblige economically weak local governments (those dependent on transfers) to shift their spending toward the priorities set by the government that pays the grant. Consequently, the structure of local priorities and spending may be deformed, and the local government’s latitude becomes smaller—which provokes a centralizing tendency.

If one considers the objectives of financial transfers described at the start of the chapter and the three pairs of criteria discussed (use, method of calculation, availability of funds), one can have four typical situations that correspond in fact to expected results (the last line of Schema 5):

1/ **Incentives** to produce or increase the quality and quantity of local public services can be generated with specific and lump-sum grants:
   - specific, to link use of the sums received to the function to encourage; and
   - lump-sum (non-matching), because transfer amounts must not depend on the spending that the recipient local government is spontaneously willing to devote to this function since this spending is insufficient.

2/ **Compensation** for the needs and cost of the local public services that must comply with the standards on merit goods mandated by the center can be provided through specific, matching, and closed-ended grants:
   - specific to guarantee that local spending is directed to the merit goods;
   - matching because the calculation of the amount depends on the gap between needs or cost compared to the standard set for the subsidized function;
   - closed-ended because closed-ended grants stimulate local spending more than open-ended grants.

3/ If the intervention is **corrective** in nature, for example in regard to the spillover of benefits derived from local public services, grants must be specific, matching and open-ended:
   - specific, that is to say linked to the function that generates the spillover;

---

20 If local spending is done based on service contracts, the level of government that pays the grant can also calculate the presumed amount of spending. In this case, one can also use specific lump-sum grants.
- matching, or proportional to the spending generated by this effect;
- open-ended so that the production of an LPG with a positive spillover is not restricted by budget limits: the extent of the externality is what is decisive.

4/ **Revenue sharing**, either to distribute financial means in a somewhat uniform tax system or to overcome a budgetary deficit, can be done through general and non-matching grants:
- general: the amount is freely available to the recipient local government;
- non-matching: the objective of resource sharing is not to steer the supply of local public services toward a specific use. The amount to share should be pre-determined and guaranteed.

5/ **Equalization** also justifies general and non-matching grants.
- general: the amount must be freely available to the recipient local government;
- non-matching: the amount is pre-determined and guaranteed.

In this case, the difference between these grants and resource sharing grants is related solely to distribution criteria: resource sharing is done in accordance with the criterion of the origin of the funds, while redistribution in the strict sense involves the financial capacity of the recipient local government. In the case of redistribution, access to transfers is sometimes subject to additional conditions such as a minimum fiscal effort or prior reorganization of how the local government manages its finances (such as applying the polluter pays principle for environmental functions).

This typology, given ex ante, must be interpreted with care because the efficiency of financial transfers, and thereby the choice of a specific form of grant, depends on the impacts, studied from the standpoint of the recipient local government or the government issuing the grant(s).

6.3. **General Principles**

The following general principles make it possible to structure and allocate budget transfers (called subsidies below) between levels of government. The proposals below are mainly drawn from a well-founded working document on the results of international comparisons of subsidy allocation criteria (World Bank, 2005).

P1. The objectives must be clear and transparent and the number of objectives for each subsidy must be minimal (one objective = one subsidy is the ideal basic rule).

P2. Subsidies must contribute appropriately to financing vertical budget gaps between the assigned functions and own revenue.

P3. Subsidies must lessen the gap between local administrations’ capacity to contribute and their needs and spending.

P4. Subsidies must preserve local budget autonomy and take into account the constraints implied by national priorities.
P5. Subsidies must support as much as possible local mobilization of resources and not penalize local governments that have succeeded in increasing their own resources.

P6. Subsidy allocation must be transparent and based on a formula that cannot be manipulated (by either the local authorities or the central government) with criteria and clear weights for each criterion.

P7. Subsidies must be stable, predictable and paid in a timely manner.

P8. Setting up performance-based subsidies that serve as incentives to improve performance can be useful.\(^{21}\)

P9. In all cases, it is appropriate to avoid covering deficits generated by poor financial practices after the fact (discontinue bail out practices).

\(^{21}\) For an overview of performance-based grants, see the document titled: *Conceptual Basis for Performance-Based Grant Systems and Selected International Experiences* by Jesper Steffensen Henrik Fredborg Larsen (2005).
7

Budget Discipline and Constraint

While public finance has for decades taken an interest in the issues of public deficits, debt and budgetary discipline—or more exactly “lack of discipline” as this is the source of economists’ concerns—to arrive at “soft” or “hard” rules of budget constraint, the same questions were only raised in the context of fiscal federalism much more recently. There are, in fact, three issues: (i) local governments’ opportunity to borrow funds; (ii) the fundamental question of hard or soft budget constraint at the local level; and finally (iii) when there is a national limit on deficits and debt (as in the Maastricht Treaty), the question of how “rights” to deficits and debt are distributed among levels of government. It is not uninteresting to note that recent contributions have focused primarily on the second question (Chabellard and Guihéry, 2005). Are we to understand that the answers to the other two questions “are obvious”? We doubt it.

7.1. Local Governments’ Borrowing Capacity

Local recourse to borrowing can be justified for two reasons, investments and cash-flow needs. Based on the “golden rule” and the “pay-as-you-use” principle of public finance, borrowing is judicious for investment and capital expenditures for it spreads the burden of investment over time through amortization (beneficiaries and payers will be the same in time horizon). It is unreasonable if it only serves to make up for insufficient funds in the operating budget (Dafflon, 1998: 187).

- First and foremost, borrowing is an efficient tool to finance investments. Combined with a strict policy of amortization based on the lifetime of the project thus financed, borrowing makes it possible to spread out the cost of an investment project whose lifetime means that future generations will benefit from it (in this case, they will participate in its financing). This is the application of the pay-as-you-use principle defended by Musgrave (1959). Among other things, financing public infrastructures through borrowing can stimulate regional development.

Of course, there is much hesitation to borrow. Swianiewicz (2004) and Dafflon and Beer-Tóth (2009) cite the following reasons for the twelve countries that have recently joined the European Union:

- financial markets do not work satisfactorily;
- commercial banks’ interest rates are too high and the risk premiums added for local governments are prohibitive;

---

22 Prud’Homme (1995), Dafflon (1995b) and Ter-Minassian and Craig (1997) were among the first to raise this question. Let us also cite the following three works that examined this situation: Dafflon (ed.), 2002, Rodden et al. (eds.), 2003, and Swianiewicz, P. (ed), 2004.
- debt is conditional on creditors (the credit market when it exists) being able to attribute a credit risk to local governments (which they cannot do when creditors rightly or wrongly believe that this risk will ultimately be borne by the central government given de facto financial solidarity between local governments);
- the borrowing and audit rules to which local governments are subject are too strict (see below on the debate over budget discipline);
- local officials think that debt is a sign of financial instability, poor management (not distinguishing between debt due to investment and debt due to cumulated deficits), or an overly risky form of financing, or that debt transfers an unfair burden on future generations (again, referencing the wrong counterpart: cumulated deficits rather than investments);
- local governments’ role is limited to day-to-day affairs (or, inversely: central government ministries have control over local investments);
- it is better to use (request, demand would be more appropriate verbs here) financial transfers from high levels of government or even NGOs and international aid agencies as substitutes for local debt;
- future financial resources are difficult to predict or even irregular and local governments are unable to guarantee that debt servicing and the operation costs resulting from the investments will be covered over long periods; and
- finally, local governments have neither sufficient knowledge of financial techniques nor the managerial capacity to manage local public debt above a certain threshold.

These obstacles can exist alone or in combination. They differ from one country to the next, but the result is that even when borrowing is entirely appropriate to finance investments, local governments prefer to use other forms of financing. Investment grants paid by the center, the European Union or other donors, public-private partnerships, and even purely and simply renouncing local public responsibility through privatization are idealized—often mistakenly—as much better alternatives. Dafflon and Beer-Tóth (2009), however, defend the thesis that these arguments are excuses and that the countries concerned should (and could) take on real responsibility in the “investment = loan” duo.

- Once these (false) arguments against borrowing have been eliminated, one must see that there are budget constraints on debt that are very real. The pay-as-you-use principle of investment and borrowing demands that amortization be real, that is to say financial and manifest by the real repayment of the loan over the lifetime of the project.\textsuperscript{23} The local government that carries the loan must also be able to pay the interest on it. But, in addition to servicing the debt, two categories of costs then emerge and must be included in the operating budget: the operating costs of running and maintaining the new infrastructure, and the functional expenses for the new (or expanded) service that the investment allows. In other words, if a local authority in a

\textsuperscript{23} On amortization methods and their impact on budgets, resource allocation and equity, see Dafflon, 1998: 179-197.
developing country invests to build a health care center destined to provide health care to the resident population, future budgets must provide for: loan interest and amortization + technical maintenance costs for the health care center (water, energy, building maintenance, maintenance staff) + the operating costs for the new function (health care staff salaries, basic medications, etc.). The accumulation of these four cost categories (interest, amortization, operating expenses, functional spending) is, in reality, the only real obstacle to financing new infrastructures through borrowing: there is no point in borrowing to obtain a new infrastructure if it cannot be used afterward because of insufficient funds. The simplest formula is (Dafflon and Beertóth, 2009):

\[
\Delta I = \Delta B + F
\]

\[
\Delta B = \frac{S - [(M + E) - (R + O)] \times 100}{i + d}
\]

where

\(\Delta I\) new investment

\(\Delta B\) borrowing capacity for a new investment

\(F\) other outside sources of possible funding for the investment

\(S\) budget / current account balance

\(M\) operating costs of the new investment

\(E\) future annual functional spending for the new function

\(R\) possible revenue from the new investment or the function that it enables

\(O\) operational grants or financial transfers from other governments

\(i\) passive interest rate on the loan, in %

\(d\) loan amortization rate, in %

This formula has the advantage of highlighting both the budget orthodoxy of the “golden rule” governing operating budget balance and the “investment = loan” duo, and also financing from outside sources, notably financial transfers from a higher level of government—which is particularly important for local governments in developing countries.

1/ Without the symbols \(F\) and \(O\) for outside financing, the local government would have to bear the weight of the new investment alone. The formula therefore calculates its borrowing capacity and, thereby, its investment capacity.

2/ Without \(F\) and \(O\) again, the balance \(S\) of the operating account would have to support all the new expenses generated by the investment. This is the concept of budget accountability in action. The local government must be able to support the costly future consequences of its spending decisions.

3/ This logic also holds for outside financing of local investments: if the new infrastructure must be financed by transfers from a higher level of government \((F)\), one must ensure that the recipient local government—with its own future resources,
transferred resources (O), or a combination of the two—will be able to utilize the investment and make its utilization and utility sustainable. Here, we have a direct link with the issue of financial transfers discussed in the previous chapter. Let us add that the local government is not the only one with budget accountability. When the central government demands or supports (with F) a local investment, it must also ensure that the local government in question will have adequate resources (either because resource sharing between levels of government gives it the resources to do so, or because it will receive operating grants or equalization grants (O)).

- In regard to short-term cash flow, borrowing facilities can overcome the delay that often exists between spending and the collection of tax revenues. We shall emit only one reservation, but it is important. Here, the aim is to correct annual short-term delays between tax revenues, transfers owed and operating expenses but not to cover an operating deficit. This warning is all the more imperative when deficits are recurrent.

7.2. Budget Constraint

Consideration of budgetary indiscipline, through the notion of hard budget constraint (the adjectives “hard” and “severe” are equivalent), is recent in fiscal federalism. According to Chabellard and Guihéry (2005: 339), this is currently a growing line of research for which there is no fixed theoretical framework and that is not treated in a uniform manner. The research must identify the characteristics of the inter-governmental financial relations behind local governments’ strategies and, in particular, explain the central government’s inability to credibly commit to not granting financial support after the fact to local governments in financial difficulty.

Before going further on this issue, let us give readers several definitions for use below:

- A local government’s **budget accountability** is respected when it supports the local tax consequences of its spending decisions directly. The current budget is balanced through the “golden rule.” Financing is neither external through transfers received nor delayed in time in transferring the tax burden on future generations via debt.
- In principle, **budget discipline** has the same characteristics as budget accountability except that it is imposed from the outside by a constitutional or legal rule to follow. The Maastricht criteria (deficit < 3% of GDP and debt < 60% of GDP) fall under this aspect.
- **Hard, strict, strong, severe** (other analogous adjectives are possible) **budget constraint** results from a financial rule nearing, or even imposing, the “golden rule” of a balanced budget and balanced public accounts.
- **Soft, loose, weak, gentle budget constraint** equals—or almost—a lack of budget rules: either budgets do not need to be balanced and debt can be used to rectify the budget.
- If a higher level of government intervenes ex post to cover and take over this deficit, it creates a situation of **bail out** in public finance jargon.
Budget decentralization runs up against a major problem in many developed countries and even more in developing and transition countries: the “weakness” of local governments’ budget constraint (soft budget constraint). This problem was initially identified by Kornai (1979, 1986) in the context of businesses under state control in the former planned economies of Eastern Europe. The same type of problem arises in the framework of public sector decentralization when the government has limited control over local governments’ spending and borrowing capacity but for one reason or another maintains a say over how they “conduct their affairs.”

A **soft budget constraint** problem arises more specifically when local governments know (even if they are not certain) that the central government will grant them additional funds (transfers) in the case of financial difficulty. Local governments then take into the higher government’s difficulty committing to not bailing them out (problem of ex ante commitment and moral hazard) and will pursue strategies that may increase their financial difficulties. From the theoretical standpoint, the problem of soft budget constraint is generally modeled as a sequence. In the first stage, the central government announces the amount of budget transfers that it is ready to grant local governments. In the second stage, the latter determine their level of spending and borrowing. In the last stage in the sequence, the central government can decide whether or not to help a local government in financial difficulty. Obviously, everything depends on the central government’s incentive to intervene after the fact. In short, two conditions must be met for there to be soft constraint:

(i) the local government must have an interest in pursuing strategies to oblige the higher-level government to bail it out or at least help it financially; and

(ii) the central government must have an interest in deviating from the policy it announced in stage one and bailing out the local government in stage three.

Technically, under **hard budget constraint**, the efficient level of public spending and debt is attained when the marginal benefit from one additional expenditure or one additional loan is equal to the marginal cost. When there is soft constraint, the marginal benefit exceeds the marginal cost perceived by the local government that expects the central government to cover part of the cost (in this case, part of the cost is shifted to national taxpayers—the common pool effect).

As we can see, the crucial question is under what conditions will the central government have an interest in not sticking to its original position, enabling the regional governments to anticipate a bail out?  

**Vertical Gap**

This problem is likely to arise when there is a vertical gap between local governments’ revenues and their necessary spending. As we have already seen, their own revenue (over which local governments have, what is more, very often only limited discretionary power—which intensifies, as we shall see, the problem of soft budget constraint) is very frequently less than spending so that part of local governments’ financing comes from budget transfers from the central government. The issue is less sensitive when local governments have enough of their own revenue to finance their

---

24 See Vignault, 2005, for a review of writings on the subject.
spending and service their debt. In this regard, readers will recall that it is more difficult to make local governments with limited resources responsible for financial difficulties. Among other things, creditors generally consider, in this case, that there is a form of solidarity between the central or federal government and the regional or local governments. Very close to home (there is no need to go to developing countries), we can see that in Germany the Länder’s debts receive a AAA rating even though some have had financial difficulties.

✓ Fiscal Autonomy and Own Resources
In some federal states (Germany is a good example again), the local governments have access to a diverse number of tax bases but have very little latitude to chose rates or bases. A large share of tax revenue comes from “revenue sharing” (see definition above) so that, here too, creditors and voters alike (the latter, let us not forget, are the first concerned) can think that local governments are not fully responsible for financial crises.

✓ Budget Transfers
Let us emphasize outright that even when there is a pronounced vertical gap between own resources and local spending, allocating budget transfers according to a transparent rule should reduce the possibility of soft budget constraint. This is the case when the transfers are determined by a clear formula—based, for example, on the number of beneficiaries (number of needy people or school children, etc.)—that ties the central government’s hands and reduced the possibility of discretionary transfers. Inversely, if allocation methods are poorly defined, then local governments may believe that they could receive additional transfers from the central government. This argument cuts both ways, however. Indeed, the best intentions in the world and a clear rule, for instance redistributing spending needs throughout the territory (with the premise that everyone must have the same quality, or the same access to, public services), can be the source of unwanted incentives and cause local governments to expect a bail out from the central government after the fact as they know its objective is to preserve the principle of territorial equity.

✓ Assignment of Responsibilities to Levels of Government
We have already emphasized that, very often, local governments are required to provide a certain number of services for the central government, services over which they have very little control (agency role). This is the case when the central government sets standards so that certain basic services are available in the country. Obviously, this exacerbates the risk of soft budget constraint. In an analogous fashion, the risk is also exacerbated when several levels of government share responsibility for providing certain public goods and services. Voters are no longer able to assign responsibility to one or the other level. We shall also insist on the fact that a large local government will obtain a bail out more easily than a small one in virtue of the “too big to fail” principle (Wildasin, 1997).

✓ Political Considerations
Political considerations are never very far from the issue of soft budget constraint. Indeed, local governments are, in a large number of countries, represented in the
upper chambers of parliament so that the central government’s desire to limit their autonomy has very little chance of becoming reality, especially in federations. Similarly, trading votes can enable a struggling local government to receive support from one or more other local governments. The central government can also have an interest in helping a local government when the majority of votes there are for the party in power or when that part of the country is a breeding ground for the party in power’s supporters. Obviously, this can also be said of other ethical or religious considerations.

**Reputation and Information**

The reputation factor is obviously very important for the central government because the local governments will anticipate its future behavior based on how it has behaved in the past. The question of the auditing and transparency of regional and local entities’ accounts is obviously very important because if budgets and accounts lack clarity local governments may blame their financial difficulties on outside factors.

7.3. **Assignment of “Rights” to Deficit and Debt to Levels of Government**

The question of sharing “rights” to deficit and debt among the various levels of government, an issue raised with intensity since the introduction of the Maastricht criteria, is addressed only a little by the theory of fiscal federalism. This topic is seen as minor and part of the practical and administrative dimension of budgets. But is it minor? Of course not.

To illustrate the nature of the problems, let us imagine a three-level system of government similar to the Swiss government. What would happen to the national budget policy if the Maastricht criteria were added to the existing constitutional provisions? In a word, who decides how to divide the 3% of GDP threshold for current public deficit and what distribution key will be used? 25 One can be certain that the classic functions of public service—allocation, distribution and stabilization—will undergo considerable change according to the institutional solution adopted concretely (which will most likely depend on the balance of power between levels of government).

From the theoretical standpoint, the specter of possible scenarios is easily spotted:

- either the 3% of GDP measure applies entirely to the public deficit at a given level (in Switzerland, the Confederation, cantons or communes, and to social insurance, in a mutually exclusive manner);
- or the quota is divided between the (three) levels of government according to a key to be written into the constitution and defined after political debate on the institutional level.

---

25 Only the question of the 3% distribution of tolerated deficit is addressed here. Answering this question does not yet equal determining *ipso facto* the distribution of total debt tolerated at 60% of GDP. Indeed, the two questions would only be answered simultaneously if one were initially to consider zero debt. However, as each level of government’s stock of debt when the deficit < 3% of GDP rule is applied is not zero, the question of the initial right to a stock of debt still remains—and does so in a logically separate manner. This aspect is not covered here, but is based on a relatively similar set of issues.
The problems rise when one moves from theory to practice, especially in light of the large number of public sector powers, the assignment of these powers (which can be revised periodically) to levels of government, and the multitude of existing legal bases on which to set the distribution key for the debt ratio tolerated by the Treaty and the complexity of the negotiating process involved. Let us examine, in order, the two cases evoked above.

The Right to Deficit Assigned to Only One Level of Government

If one were to simply assign the amount of deficit tolerated by the Maastricht standards to the central government’s (the Confederation’s) budget, the lower levels (read: the cantons and communes) would logically no longer have the possibility of taking out loans to manage decentralized public finances. This state of affairs could only emasculate the local administrations’ “choice” role by pushing to the extreme their function of simple executing agency for the policies decided (solely) by centripetal forces and by sacrificing their power to offer local public goods in response to their own residents’ preferences (Smith, 1996: 277-278). The probable increase in the amount of financial transfers from the center would, then, be the most tangible expression of balancing the budget at the price of lost autonomy.

If, however, the prerogative of contracting debt applies exclusively to regional governments (the cantons, in Switzerland), the higher level of government would no longer be able to act in financial markets and would consequently have to repay (spread over time) the bonds issued in the past. It is obvious that this situation would seriously endanger the public functions that only the central government is capable of fulfilling efficiently (the canonic example is national defense) unless one admit that the central state match investment spending to its net margin of annual self-funding—which would deprive it of any situational leverage or would even reinforce pro-cyclic adjustments. In “good times” (fiscally profitable periods), the central government would, therefore, make up for the delays accumulated during “bad times.” We can obviously imagine an intermediary scenario, such as what Belgium did, in which the right to deficit is of course shared but by the federal government that can limit a community’s or a region’s borrowing capacity (and thereby its deficit), and can in fine exercise veto power (Deschamps, 1994: 131-132).

Furthermore, when there are several government entities on the same hierarchical level (regions or local governments), the problem of the internal distribution of the tolerated amount of the nation-state’s public deficit arises. Let us illustrate this last point with the help of the following paragraph, which attempts to address this phenomenon from the angle of “vertical” relations.

Tolerated Deficit Shared Between Several Governmental Bodies

When several levels of government must share the public debt within the 3% of GDP limit, assigning tolerated deficit quotas becomes an elaborate exercise in political economics, given both the heterogeneity of existing budget rules and the strategic behaviors dictated by institutional interests (Dafflon, 1997: 106-110). For the moment, let us consider only the first factor; we shall examine the second afterward. Taking the Swiss situation as a model, sharing the 3% of GDP annual deficit limit between the
Confederation, the 26 cantons (and semi-cantons), and the 2,551 (2011) communes implies recourse to (1+26+26=) 53 distinct regulations, the budget parameters of which must be made comparable to each other for operational reasons. This calculation needs explanation. Three legislative categories exist alongside each other in fact: the federal law on the Confederation’s finances, the twenty-six financial laws governing cantonal budget management (according to the “one canton, one law” principle of autonomy), and the twenty-six cantonal laws on communes. In total, if one wants to evaluate the effect that applying the Maastricht criteria to federal finance in Switzerland would have, one would therefore have to include in the same area of reference 53 legal texts that do not necessarily have the same content. The following scenarios are possible:

(a) decision power belongs exclusively to the federal government that decides for itself (1x), the cantons (1x) and the communes (1x) = 3 laws or regulations;

(b) it belongs to the federal government that decides for itself (1x) and the cantons (for both levels cantons + communes) (1x), leaving each canton to assign its individual quota to itself and its communes (26x) = 28 laws; or

(c) the scenario described above, which contains 53 laws.

Scenarios (a) and (b) obviously raise the question of the cantons’ autonomy and seem difficult to implement a priori. The objective difficulty of an institutional deficit-sharing solution is the rigidity of such solutions that, hypothetically, also freeze the assignment of powers and revenues.

✓ **Strategic Behaviors**

Breaking with the solidarity principle, we can imagine that the central government shifts the weight of balancing the budget to the local governments. In a three-level system of government, the degree of freedom at the third level is nil. Here, the Länder (cantons) are placed in a privileged situation because they are in an intermediary position. In short, they can shift the burden to the communes in two different ways:

- The cantons have the ability to limit the commune’s “choice” role by obliging them to take on only the role of simple executing agency for functions decided by the higher level of government (*cf.* above).

- The communes can be placed under supervision fairly easily, especially since they do not form a formidable pressure group and the supervision can be indirect. Granting a deficit or debt limit that is a priori generous is thus a response to a new assignment of functions that puts local governments in charge of providing certain services or shifts provision of certain services from the center to the cantons (decentralizing only the production function and not supply).
Conclusion

The purpose of the preceding discussions was to provide an interpretation guide to the theory of fiscal federalism. We have attempted to show that the notion of decentralization was contingent upon the history and institutions of each country. It is therefore illusory to want to apply the same schema of decentralization in all countries. This is all the more true for developing countries because many of the hypotheses formulated by the theory of fiscal federalism are inoperable in these countries. Consequently, the only possible method, given the current state of knowledge, is for practitioners to ask the right questions based on a methodology derived from both lessons from economic analysis (and the related disciplines) and experience acquired in the field. Let us summarize the essential points by way of conclusion:

The message in the second chapter is that there are undeniable advantages to well-conducted decentralization (in both meanings, “delegation” and “devolution”). Local public services can be tailored to the preferences of the residents in each local government unit in accordance with the demands of good resource allocation and avoidance of waste under “national standards”. The risks involved in innovating and trying new methods are contained—but innovations may still be imitated afterward if they “pay off”. The relationship between the services provided and their fiscal cost (the taxes to pay) obliges local politicians to be attentive to local voters’ wants and to what local taxpayers are willing to pay. In this way, the circles of decision-makers, payers and beneficiaries are brought closer together.

Chapter three is summarized in the decentralization matrix and contains three strong messages. (i) The theory of fiscal federalism does not provide ready-made answers to the problems of any given federal state nor do national practices, but we can learn from both (Blindenbacher and Koller, 2003). (ii) Before discussing the assignment of functions to levels of government, it is indispensable that the stakeholders in the decentralization process harmonize the vocabulary, expressions and concepts used, and that they discuss the criteria to use and the possible institutional organization of decentralization, from the communes to the central government. (iii) The decentralization matrix, used for each function, highlights the advantages and disadvantages of (de)centralization from the allocative, redistributive, macroeconomic and budgetary standpoints. It allows one to ensure that choices are coherent, but leaves politicians in charge of weighing factors when objectives diverge.

Chapter four compares the budget autonomy and accountability that give local governments access to taxes with the difficulties of managing a decentralized tax (in terms of power to determine and levy taxes and handle disagreements) —and tax competition as an external constraint on local budget policy. In short, it is not enough for
federated entities to have a constitutional right to levy taxes; they must also have the political will to do so, accept responsibility for doing so, and withstand tax competition.

Based on the decentralized budget, chapter five looks at the concepts of budgetary and financial autonomy. Definitions are not simple, and each adds nuances. Above all, one must reject false ideas: being “fiscally sovereign” in one’s local capacity to levy a tax does not *ipso facto* grant one autonomy in budget decisions. A small commune can be wealthy from taxes but too small and incapable of providing alone the functions and services that one could expect from a local entity in a given country. Local autonomy will only ever be marginal: the analytic interest is to measure latitude to make decisions and analyze how it is used.

In the field of financial transfers, national situations can be summarized by the expression “anything, any way.” Countless formulas have been invented. Chapter six addresses them. It provides nothing new, simply an interpretation guide that allows one to bring order, a little transparency and, we hope, coherence to the jungle of one-off formulas.

Chapter seven, devoted to budget constraint and borrowing, can be called painful: ultimately, it is indeed their “pocketbooks” that determine the extent of the powers really exercised by decentralized entities. Our preference goes to compliance with the golden rule when balancing local operating budgets. The concept of budget accountability is what confers rights (notably the right to borrow to finance investments) but it also confers duties—paying passive interest and amortizing the loans, and ensuring that future financial resources will be sufficient to make full use of the new investment. This is the meaning given to budget accountability. The other component, macroeconomic this time, is budget discipline: how can one combine local budget accountability with binding rules on deficits and debt—in the sense of the Maastricht rules, for example.

We shall finish by insisting that this contribution is far from an exhaustive tour of the theme of fiscal federalism and decentralization. Many questions deserve more in-depth treatment, for instance those on cooperation between or mergers of communes in response to excessive political fragmentation, redistribution as a model of solidarity (vertical, horizontal?) between levels and decentralized entities, transfers between levels of government (such as, for example, deciding between non-matching grants written into the constitution as an alternative to local taxes), or even the introduction of participatory democracy mechanisms (right of initiative, popular referendum, referendum by federated entities) as a substitute for (supplement to?) the rules of budget discipline imposed by the center.
Bibliography


**Back Cover**

During the past two decades, decentralization issues have generated keen interest in a large number of countries. This notion is, however, difficult to define as it refers to a very wide range of institutional arrangements. The purpose of this document is to provide an interpretation guide for decentralization-centralization processes based on the theory of fiscal federalism. The idea is not to identify one decentralization schema for all countries as decentralization is contingent upon the history and institutions of each country. This is all the more true for developing countries because many of the hypotheses formulated by the theory of fiscal federalism are inoperable in these countries. The goal is, rather, to help practitioners ask the right questions based on a methodology derived from both the lessons of economic analysis (fiscal federalism and the related disciplines) and experience acquired in the field.