

International Studies Program

Working Paper 04-01
January 2004

Subnational Government Structure and Intergovernmental Fiscal Relations: An Overlooked Dimension of Decentralization

Jamie Boex
Jorge Martinez-Vázquez
Andrey Timofeev



Georgia State
University

Andrew Young
School of Policy Studies



Subnational Government Structure and Intergovernmental Fiscal Relations: An Overlooked Dimension of Decentralization

Working Paper 04-01

Jamie Boex
Jorge Martinez-Vázquez
Andrey Timofeev

January 2004

International Studies Program
Andrew Young School of Policy Studies
Georgia State University
Atlanta, Georgia 30303
United States of America

Phone: (404) 651-1144
Fax: (404) 651-3996
Email: ispaysps@gsu.edu
Internet: <http://isp-aysps.gsu.edu>

Copyright 2001, the Andrew Young School of Policy Studies, Georgia State University. No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means without prior written permission from the copyright owner.

International Studies Program Andrew Young School of Policy Studies

The Andrew Young School of Policy Studies was established at Georgia State University with the objective of promoting excellence in the design, implementation, and evaluation of public policy. In addition to two academic departments (economics and public administration), the Andrew Young School houses seven leading research centers and policy programs, including the International Studies Program.

The mission of the International Studies Program is to provide academic and professional training, applied research, and technical assistance in support of sound public policy and sustainable economic growth in developing and transitional economies.

The International Studies Program at the Andrew Young School of Policy Studies is recognized worldwide for its efforts in support of economic and public policy reforms through technical assistance and training around the world. This reputation has been built serving a diverse client base, including the World Bank, the U.S. Agency for International Development (USAID), the United Nations Development Programme (UNDP), finance ministries, government organizations, legislative bodies and private sector institutions.

The success of the International Studies Program reflects the breadth and depth of the in-house technical expertise that the International Studies Program can draw upon. The Andrew Young School's faculty are leading experts in economics and public policy and have authored books, published in major academic and technical journals, and have extensive experience in designing and implementing technical assistance and training programs. Andrew Young School faculty have been active in policy reform in over 40 countries around the world. Our technical assistance strategy is not to merely provide technical prescriptions for policy reform, but to engage in a collaborative effort with the host government and donor agency to identify and analyze the issues at hand, arrive at policy solutions and implement reforms.

The International Studies Program specializes in four broad policy areas:

- **Fiscal policy**, including tax reforms, public expenditure reviews, tax administration reform
- **Fiscal decentralization**, including fiscal decentralization reforms, design of intergovernmental transfer systems, urban government finance
- **Budgeting and fiscal management**, including local government budgeting, performance-based budgeting, capital budgeting, multi-year budgeting
- **Economic analysis and revenue forecasting**, including micro-simulation, time series forecasting,

For more information about our technical assistance activities and training programs, please visit our website at <http://isp-aysps.gsu.edu> or contact us by email at ispaysps@gsu.edu.

Subnational Government Structure and Intergovernmental Fiscal Relations: An Overlooked Dimension of Decentralization

Jamie Boex, Jorge Martinez-Vazquez and Andrey Timofeev
Andrew Young School of Policy Studies, Georgia State University

1. Introduction

Many countries around the world are currently engaged in one way or another in policy debates or reforms of their system of intergovernmental fiscal relations.¹ While the policy dynamics vary greatly from country to country, the focus of such reform efforts can generally be categorized into one of the four main dimensions of intergovernmental fiscal relations, namely the assignment of functional responsibilities (expenditure assignments), the assignment of revenue sources, the provision of intergovernmental fiscal transfers, and the regulatory framework for subnational borrowing and debt. Based on the mantra that “finance should follow function,” the conceptual starting point of fiscal decentralization reforms is generally held to be the assignment of expenditure responsibilities, which seeks to address the question what functions and expenditure responsibilities should be assigned to each level of government (Bahl 1999). However, a preceding question that is often overlooked pertains to the structure of subnational governments: how should the government sector be structured in order to ensure that fiscal decentralization reforms will result in a more efficient allocation of public resources?

¹ Since such reforms of intergovernmental fiscal relations are typically aimed at providing government services in a more decentralized manner, these reforms are more commonly referred to as fiscal decentralization reforms.

The matter of the structure of subnational government –how many government levels should exist and the determination at which levels of geographical disaggregation different tiers or levels of government should exist- is typically left to historical accident and has gone essentially unquestioned in the fiscal decentralization literature. In actuality, however, a sound structure of subnational government levels or tiers is crucial in assuring a sound assignment of expenditure responsibilities. Without an appropriate subnational government structure, it would be impossible to realize the efficient delivery of subnational government services, achieve correspondence in local taxation, and ultimately accomplish the benefits of decentralization reforms.

Subnational government structure: a relevant policy issue world-wide

The question of subnational government structure is becoming an increasingly important issue across countries pursuing decentralization reforms around the world, throughout different continents and irrespective of the stage of development of the economy in question. For instance, in Russia's gradual transition towards a new federalism, the Russian government is not only contemplating how it should structure the relationship between the federal government and the regional governments, but also to define the intergovernmental relations below the regional level by delineating the relationship between county-like *rayon* governments and (potentially sub-*rayon* level) municipal governments (see the appendix). In many transition economies, such as the Czech Republic and Hungary, the introduction of local self-government in the 1990s led to excessive fragmentation of local authorities due to thousands of small villages claiming the status of an autonomous municipality, in part in reaction to the forced consolidations and deprivation of services and development under the previous, highly centralized system (Davey, 2002).

In a similar vein, the need to reconsider and potentially reform local government structures is a common theme in many developing countries. In many countries in Sub-Saharan Africa (such as Uganda and Tanzania), policy questions are being raised regarding the structure, respective roles and the division of responsibilities between district level governments and sub-district governments such as wards and villages. In response to what was perceived to be excessive fragmentation of the local government structure following the introduction of South Africa's 1994 Constitution, a restructuring of municipal jurisdictions in 1998 brought down the total number of municipalities from 843 to 284. The reform was aimed at achieving viably-sized local jurisdictions, the equalization of revenue capacity, and achieving racial integration. Likewise, policy makers in Nepal are debating whether districts and villages in their country are too small to assure the efficient and effective delivery of local government services, or whether consolidation of districts and villages would assure a more appropriate structure. Bangladesh –possibly one of the world's most centralized countries- is in ongoing process of reconsidering its multi-tiered structure of subnational governments. In particular, Bangladesh is considering reintroducing local elections at the Upazila Parishad level (currently the lowest deconcentrated tier of central government) and expanding the Upazila's role in supervision and delivery of government services. Bangladesh is further

grappling with the question whether there ought to be an elected village level of government.

Well-established democratic and industrialized countries are not immune from questions of subnational government structure. In Western Europe, unitary states such as the United Kingdom as well as highly centralized France have moved in recent years to provide increased (fiscal) autonomy to particular regions; notably Corsica in France and Scotland and Wales in the United Kingdom (BBC News 2003; Young 2001). In 1993, Belgium took the drastic step of moving towards a federal government structure, with important implications for intergovernmental fiscal relations. In Denmark, a government commission appointed in 2003 to study the country's government structure indicated that a large number of the current administrative units are too small and recommended further consolidation of municipalities and rural district governments (Danish Ministry of Interior and Health 2004). In Canada, in addition to the perennial challenge posed by Quebec, the policy question is periodically raised whether the federal government should establish direct fiscal relations with local governments, or whether local governments should continue to fall exclusively within the realm of provincial governments. In short, although the issue emerges in a number of different forms, the question of subnational government structure is highly relevant in countries around the world.

Three dimensions that define subnational government structure

While decentralized governance may have (non-economic) benefits in assuring a balanced governance structure and political stability, the most frequent purpose of decentralization reform focuses on the fiscal dimension: fiscal decentralization reforms are often pursued to enhance the equity and efficiency of public service delivery. In order for these economic benefits from decentralization to materialize, government has to be "closer to the people": smaller subnational government jurisdictions would allow the public sector to better match government services with local demands and needs, and thus to achieve improved allocative efficiency. At the same time, economic forces require that subnational jurisdictions do not become excessively small, since larger jurisdictions are often able to deliver services more efficiently due to scale economies. In order to assure that government services are produced in a technically efficient (cost-minimizing) manner, the assignment of expenditure responsibilities should take place in such a way that smaller subnational governments are not assigned the responsibility to provide services which are subject to scale economies. In addition, subnational jurisdictions should be sufficiently large so that subnational government services are not affected by interjurisdictional spillovers (externalities). The key question then becomes, based on economic principles, what is the optimal size of subnational governments, and what is the associated optimal structure of subnational governments?

Although dominant features of a country's subnational government structure include the typical size of a local government jurisdiction at each government level, and the number of tiers or levels of subnational government that exist, in fact a series of related features

combine to define a country's subnational government structure in the context of a system of intergovernmental fiscal relations. As such, the core questions to be answered are:

1. What is the most efficient size for local government jurisdictions for the delivery of local government services?
2. How many levels or tiers of government are there or should there be in a country, and at which level(s) of disaggregation should these levels be established?
3. A third dimension of subnational government structure regards the horizontal and vertical relationships between levels of government and government units. This set of policy questions considers how subnational government jurisdictions are formed and governed, and how subnational governments relate to the central government and each other.

Structure of the paper

The purpose of this paper is to explore what guidance economic theory and international practices offer policy makers in considering the issue of subnational government structure in the context of fiscal decentralization reforms, particularly in developing and transition economies. To this effect, the remainder of this paper is structured as follows. First, Section 2 reviews the conceptual and theoretical foundations that guide the economic determination of an optimal size of local governments, while Section 3 considers the application of these concepts and theories to subnational government structure in the real world. Next, Section 4 addresses the second dimension of subnational government structure: what factors determine the optimal number of subnational government levels? Section 5 addresses the vertical and horizontal relations between different government levels. Finally, we consider the applicability of the conceptual framework by making a number of observations about international practices and experiences in structuring subnational government relations.

2. Conceptual foundations of optimal subnational government structure: the optimal size of local governments

We noted at the outset of this study that the structure of subnational government in many developing and transition countries is more the result of historical accident than sound economic design. Naturally, this statement clearly implies that economic theory can actually provide some degree of guidance in answering the question: what is the optimal size of local government jurisdictions, and what is the resulting optimal structure of local governments?

In fact, economic theory can be applied to look at this question from a number of different angles. An appropriate starting point for the discussion on local government structure might be to consider subnational governments simply as provider of goods and services similar to business firms in competitive industries. In drawing analogies with private service providers, we look at optimal jurisdiction size from three different angles. First, we approach the optimal structure of the government sector from an economic organization perspective commonly used by economists in the analysis of the organizational behavior of firms. Importantly, organizational theory suggests that an excessively centralized system (whether business or government) is detrimental to the efficiency with which goods are produced. Second, once we have established that a decentralized government structure with more than one level of government is deemed desirable based on efficiency arguments, we consider the role of subnational governments in the public sector and the nature of local government services. This leads us to conclude that many local government goods are not truly “public goods” but in fact are “club goods.” This realization has important implications for the optimal size of local government jurisdictions. A third (complementary) line of thought considers the delivery of public goods and service across space based on a set of model’s known as “location theory.”

2.1 Arguments for decentralized government structure: organizational theory and subsidiarity

Organizational theory

A first useful line of inquiry in considering local government structure is whether we can establish whether economic theory supports the optimality of a decentralized government structure in the first place, or whether the available evidence suggests that a centrally organized public sector would indeed be more efficient. We would argue that organization theory –and the application of this theory to the private sector- indeed suggest that a decentralized system of government is likely to be more efficient than a highly centralized government structure. Organizational theory is the part of the industrial organization literature that considers how the optimal size of organizations is affected by the cost of managing information and contracts alternatively within a firm and between firms. Since the firm is the basic agent in market-based industrial sectors, much of this literature (known as “*theory of the firm*”) analyzes (changes in) the size of business entities, as firms seek to optimize their efficiency in a competitive market environment.

Theory of the firm seeks to address the observation that most competitive markets are neither vertically nor horizontally centralized or dominated by a single firm. Instead, production in virtually all sectors of the economy takes place in a highly decentralized manner, with different firms involved in different vertical stages of the production process as well as significant horizontal competition at the different stages of production. This raises the question that, if large corporations are able to produce more efficiently

than small firms due to scale economies and certain other scale-related advantages, then why do we not observe a global economy dominated by a handful of vertically and horizontally integrated business firms? This question is in many ways analogous to the question why a single centralized government organization would not be the most effective government structure. As we discuss momentarily, many of the observations made by the industrial organization literature carry over to the government sector as well.

The theory of the firm considers that at the core of its activities, business firms acquire natural resources and intermediary inputs and combine it with labor and entrepreneurship in order to produce outputs that are sold to final consumers or other firms. Depending on the nature of the production process, small firms are typically unable to specialize in all stages of the production process and often end up contracting for intermediary inputs rather than controlling the entire production process internally. However, as the firm's scale of output grows, it becomes easier for firms (or any organization) to achieve increased specialized production and economies of scale in the own production of intermediary inputs, allowing larger firms to rely less on contracting out of inputs in favor of own production.

However, there are also negative dimensions to increases in the size of firms. Most importantly, as firms gets larger it inevitably becomes harder to coordinate the activities of all the departments and stakeholders within the firm and to price internal transactions between different parts of the firm, thus making it harder to ensure that all parts of the firm are working towards the company's strategic goals in an efficient manner. As a result, for large firms, it may be increasingly difficult to make internal business decisions in the absence of the pricing signals that are provided by the market mechanism. For instance, the Human Resources Department in many large firms (or the Research and Development Department for that matter, or other such departments) are completely isolated from direct market pressures, and often end up behaving more like a large bureaucracy than an integral part of a profit-making enterprise. When business organizations become so large that large parts of the company start behaving in strategic and inefficient ways, it becomes more profitable and efficient to separate a single large business enterprise into multiple, individually operated business units. In that case, the core business would revert to relying on contracting or out-sourcing to purchase key inputs into its production process in order to maintain an efficient production process. As such,

The main lesson from theory of the firm thus is that the optimal size of a service provider depends on the trade-off between the cost savings from economies of scale (which increase as the entity's size increases) on one hand, and the inefficiencies from internal resource allocation that are associated with larger firm size on the other hand. This concept appears to be as true (if not more so) for government organizations as it is for business entities.

The subsidiarity principle

The same trade-off between cost and flexibility in allocation decisions that determines the optimal size of firms in the market sector are similarly valid for the optimal provision of local government goods: local government jurisdictions should be formed in such a way that balances the production cost of the good or service with the proximity of the government unit close to the community that it serves, allowing the unit to flexibly allocate resources consistent with local needs. In the assignment of expenditure responsibilities between different levels of government, this trade-off between scale economies on one hand, and flexibility and accountability on the other hand is reflected in the concept of “subsidiarity” or the “subsidiary principle.” Often used as a guiding principle in expenditure assignment, this principle states that the responsibility for the provision of goods and services should be placed at the lowest level of government that is able to deliver the good or service in an efficient manner. In other words, subsidiarity suggests that government goods and services should be provided at the lowest level of government compatible with the smallest “benefit area” of the service.²

For instance, the benefit area for sanitation services (household garbage collection) is clearly the local community, so in accordance with the subsidiarity principle, this service should be provided (i.e. delivered and paid for) at the local government level. Since the benefit area for air traffic control is the entire national territory, this service should be provided at the central (national) government level.

Reliance on the smallest-possible efficient benefits area in the provision of subnational government services contributes to the efficiency in the delivery of public goods and services in a number of ways:

- The smaller the distance between decision-makers and the communities they serve, the better informed the policy makers will be about the preferences and needs of the people.
- The greater the fragmentation of (efficiently-sized) subnational jurisdictions, the greater the potential for allocative efficiency due to a better match between the local mix of goods and services provided and the local variations in the preferences and needs for public services.
- Because of the closer proximity to their local government officials, local residents are better informed about the actions of their governments and they can use the quality of service delivery in neighboring jurisdictions as a performance benchmark. Given the appropriate governance mechanisms, local government

² The subsidiarity principle has various implications for the role of subnational governments in the public sector. For instance, in accordance with the subsidiarity principle, subnational governments are not well-positioned to engage in income redistribution or economic stabilization policies. For a more in-depth discussion of the subsidiarity principle and the assignment of expenditure responsibilities, see Martinez-Vazquez (1998).

- officials will be more accountable to local voters for the quality of services they provide compared to central government officials.
- Local government provision of public services at a lower government level will make taxpayers more willing to pay for public services, since the link between costs and benefits is strengthened. Efficiency in the provision of public services is enhanced if consumption benefits are linked to costs of provision via fees, service charges, or local taxes.

However, the subsidiarity principle does not argue that public goods and services should be delivered by the lowest government level capable of doing so at any cost. For instance, shifting the supply of public services with a regional benefits area to a lower government level (e.g., to local governments) is likely to result in the under-provision of services, which would be inefficient. For example, if a municipal hospital (financed by municipal taxes) were expected to provide health care services to the entire surrounding region, this hospital would likely end up providing a non-optimal level of health care. Municipal residents would be unwilling to finance health care services for residents in the wider region; as a result of the mismatch, local residents would put pressure on their elected officials to reduce spending on health care services at the hospital and look for alternative means of health care provision.

Thus, the subsidiarity principle argues that the structure of government should be determined in such a way that allows the lowest level of government to provide each good or service to the extent that it can do so efficiently. In other words, local governments should be subdivided into smaller jurisdictions as long as the jurisdiction continues to be able to provide its goods and services efficiently. This implies that the first step in identifying an optimal subnational government structure is to identify the minimum efficient scale of the services assigned to the local government level; in principle, the size of local government jurisdictions should at least be equal to or greater than this minimum efficient scale.

However, it would be naïve to define the optimal number of local governments as the total population of a state or country divided by the minimum efficient scale associated with the delivery of the most important local government service. Since we cannot ignore the existence of scale economies associated with administering a number of schools or health clinics at once rather than administering each facility separately, or the scope economies that are gained from providing several different government services through a general-purpose local government, it becomes virtually impossible to concretely identify one single local provision area that achieves this standard for all goods/services provided by a local government. Nonetheless, the subsidiarity principle is a useful conceptual tool in analyzing expenditure assignments and in guiding our understanding of the optimal structure of local government.



Box 1

Scale economies and alternative modes of producing local government goods

The existence of scale economies is often used as an argument (or a pretense) to prevent devolution of an expenditure responsibility to a lower level of government. Since subnational government units at any local level typically vary in size, it is not uncommon for opponents of decentralization reform to argue that especially the smaller local governments are not able to take on the burden of government services that display a certain degree of scale economies. Such concerns have unnecessarily hindered the implementation of decentralization reform since there are several ways in which jurisdictions (particularly smaller ones) may circumvent the negative consequences of scale economies.

The crux of the argument here is that just because local governments have been assigned the responsibility to fund and provide a certain government service, this does not necessarily mean that they have to produce this good or services themselves. There are three ways in which a local government jurisdiction can provide a good or service without producing it on its own and still retain the benefits of local policy discretion and accountability:

- Contracting with the private sector. A local government may choose to contract with a private provider to produce a certain local government service, such as garbage collection or the operation of the local market. By doing so, the benefits from fiscal decentralization (including better a match between service provision and local preferences, as well as increased local accountability) are retained, while also being able to capture the scale economies and experience brought to bear by large, specialized private service providers.
- Contracting with other local governments. Another possibility to avoid the trappings of scale economies is for two smaller jurisdictions is to collaborate on the production of a certain public good. For instance, one local jurisdiction that is too small to provide secondary schooling in a cost-effective manner on its own might contract with a neighboring local government and send its students to the neighboring district. Alternatively, the two districts might establish a covenant to jointly run a secondary school for the education of children from both jurisdictions.
- Special service districts. In certain cases, special service districts could be formed for the provision of certain local public goods. This is particularly the case in situations in which several local governments are inextricably linked by the provision of a certain local government good or service, such as a water district that covers several local government jurisdictions. Again, special districts allow a system of intergovernmental relations to resolve problems with scale economies without altogether abandoning the benefits of a decentralized system of local government.

2.2 Arguments for decentralized government structure: Club goods and Tiebout

A second line of inquiry builds upon the principle of subsidiarity and the realization that many government-provided goods, especially at the subnational level, are not “pure”

public goods. The nature of the publicly provided good and the motivation for the public sector to provide this good have important implications for subnational government and the expenditure assignment. For instance, the subsidiarity principle suggests that it would not be possible for local governments to deliver purely public goods in an efficient manner, since the benefits area of public goods virtually by definition encompasses the entire nation.³ Instead, local government-provided goods and services are often similar in nature to private goods which are (largely) excludable and (largely) rival. For instance, consider local provision of schools, small parks or play grounds, libraries, and so forth. To the extent that these goods and services are not provided by the public sector for redistributive purposes (e.g., in order to assure equal access to primary education), why would the government sector provide such goods at all? The answer to this question provides additional insights into the optimal structure of local governments.

Club goods

Economists argue that in the case of many local government goods and services, market failure does not result from the nature of the good (after all, as just noted, many goods and services that are provided by local governments are excludable and rival), but from the significant scale economies in the provision of these goods, that typically makes provision by individual consumers not feasible. A common example of a service that meets these conditions is a swimming pool; a swimming pool is both excludable and (to some degree) rival, yet most households do not have a swimming pool of their own due to significant scale economies involved in building a private swimming pool.

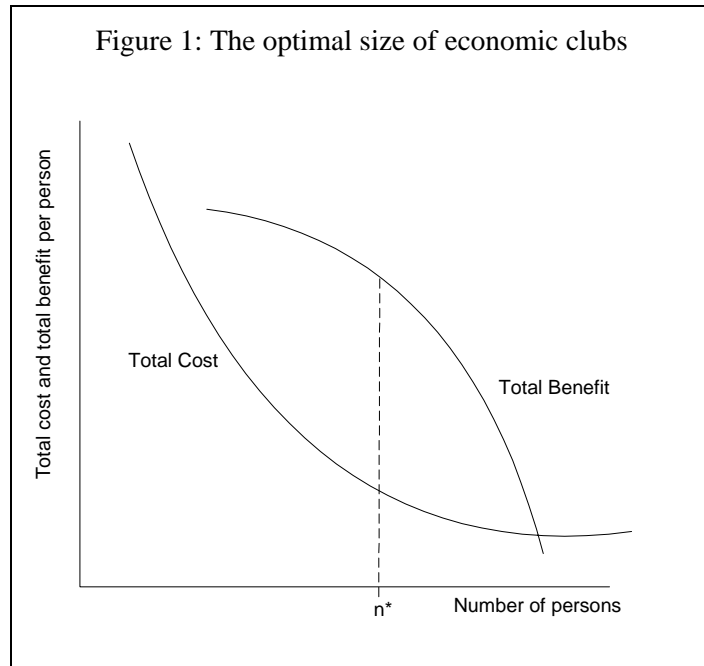
James Buchanan (1965) who first explored this issue argued that if households would get together in clubs, then these clubs could provide the relevant good to a larger number of people at a lower cost that would be affordable and acceptable to everyone in the club. This, of course, is the rationale behind community pools and public swimming pools. Buchanan thus argued that in many cases local governments simply function as clubs that assure that certain goods are provided and that collect user fees from households, or proxy user-fees in the form of local taxes. Many other local government goods and services are club goods, including local amenities such as parks and green-space, markets, child-care facilities, and primary education.

It is in principle possible to rely on private clubs to deliver these services, and we observe this to some extent, for instance, in the existence of private fitness clubs with indoor swimming pools. However, to the extent that it is costly to exclude people from certain other club good (such as a public playgrounds, outdoor pools or local parks, or even a community road or irrigation scheme), it is simply more practical to deliver the good or service by a “public club” (e.g., a local government) where all residents are members and

³ A (pure) public good is defined by economists as a good that is non-rival and non-excludable. Non-rival means that a person can enjoy the benefits of a good without reducing the benefit received by others; non-excludable means that a person cannot be excluded (or it is prohibitively expensive to exclude him/her) from enjoying the benefits of the good.

where all residents contribute to the cost of the club amenities in order to avoid free-riders.

So if we consider local governments as local clubs, within this economic theory of clubs, what determines optimal size? In other words, what would be the optimal size for local governments to provide a certain club good or service? How the optimal size of a club can be determined is illustrated in Figure 1, showing how both the cost per person varies with the club's size (number of people in the club), as well as the relationship between benefit per person and club size.



As the number of people included in the club increases, the per-person cost of providing the good or service declines as a result of scale economies. However, as the club's size increases, the benefit to each club-member decreases as more people share the facility (depending on how rival the good is and the "congestion costs" incurred). One could even argue that some of the club's members now may have more limited access to the facility as the coverage area grows (i.e., increased transportation costs). The optimal size of the club occurs where the difference between costs and benefits is maximized, which occurs in Figure 1 at n^* number of people.⁴ The theory of clubs further suggests that there is a point beyond which the total cost of a club exceeds the total benefit of the club; establishing clubs (i.e., local governments) beyond that range would thus *reduce* the welfare of its residents.

The Tiebout hypothesis

One problem of Buchanan's theory of clubs is that it is largely silent on how clubs are formed across space and how clubs are governed. If people with diverging demands for club goods locate near each other in geographic space, it would be unlikely that "clubs" (local governments) would actually form, as there would be no consensus on what level of club goods should be provided. In addition, Buchanan's theory is silent on how clubs

⁴ In terms of micro-economic optimization, this occurs where the marginal benefit of changing the club's size equals the marginal cost of doing so.

(local governments) would be governed, so that there is no assurance that local governments are managed in a way that promotes efficient provision of local government services.

Charles Tiebout (1956) built on the notion that local governments are akin to clubs formed in geographical space, and explored the efficiency properties of a system of local governments. Tiebout's hypothesis argues that the provision of local government services in a metropolitan region with a large number of local governments will result both in allocative (consumption) as well as technical (production) efficiency under certain assumptions.

The crux of the Tiebout hypothesis is that households will try to maximize their level of satisfaction, (1) by locating in the local government jurisdiction that provides the level and quality of local government services that they desire, and (2) by locating in the local government jurisdiction that provides these local government services in an efficient manner. Under the Tiebout hypothesis, local governments would provide local public goods in a cost-effective manner, and finance them through (proxy) user fees or head-taxes.

A first implication of the Tiebout hypothesis is that residents will sort themselves in "clubs" or local governments with homogenous preferences for local public goods, since all residents within a jurisdiction receive the same level of public service and pay the same price. If a household resided in a jurisdiction that did not offer its desired mix of local goods and services, the Tiebout hypothesis suggests that this household would "vote with their feet" by moving to a jurisdiction that matches its preferences. For instance, if a local government would offer an unattractive basket of public goods and services, this household would identify a jurisdiction that provides exactly the package of local government services that it desires and would relocate.

A second implication of the Tiebout hypothesis is that competition between jurisdictions prevents local governments from being inefficient or from engaging in redistributive policies (not supported by taxpayers). For instance, if a local government is excessively inefficient in the provision of local public services, not only would local government officials risk being voted out of office, but households and firms over time would also move to jurisdictions where their local tax resources would not be wasted. Similarly, if a local government would engage in redistribution, wealthier households would vote with their feet and leave the local jurisdiction, thereby eroding the local government's tax base and its ability to engage in redistributive policies.

One of the policy implications of the Tiebout hypothesis is that regions with a greater number of local governments should achieve a greater level of allocative efficiency and be more efficient in providing local public services due to (1) the increased ability of local governments to match the local needs and (2) the increased competition by local jurisdictions. A precondition, however, for such efficiencies is that households are able to

assess the relative (in)efficiency of their local government when making location decisions. As such, the Tiebout hypothesis implicitly requires transparency in the operation and administration of local governments.

2.3 *Local government structure and location theory*

While organizational theory provides an economic rationale for industries and governments alike to be organized in a decentralized manner, and the theory of club goods and Tiebout help us better understand the formation and optimal size of local clubs, neither specifically relates the provision of local government services to geographical space, which is an important dimension in the establishment of local government jurisdictions. Therefore we turn to location theory, a strand of the urban economics literature, which argues that there are external, competitive forces that help define the market structure of firms across both economic and geographical space.⁵ Again, we believe that there are important parallels between the market-based private sector of the economy and the government sector, as the basic principles of location theory can be used to better understand the optimal structure of the local government: what is the optimal size of a local governments, and how many local government units are needed to provide local government services in an optimal manner?

Location theory is based on the premise that when firms make their location decisions across space (in order to maximize their profits), firms are forced to take into account the fact that the cost that households face when making a purchase does not only consist of the price of the good or service purchased, but also of the cost that the household has to incur in order to reach the store or point of service delivery. The resulting location decisions of firms, the number of firms in a region, and the size of the market area that each firm serves (which jointly give rise to the overall market structure) can best be explained through an example.

An introductory example of location theory

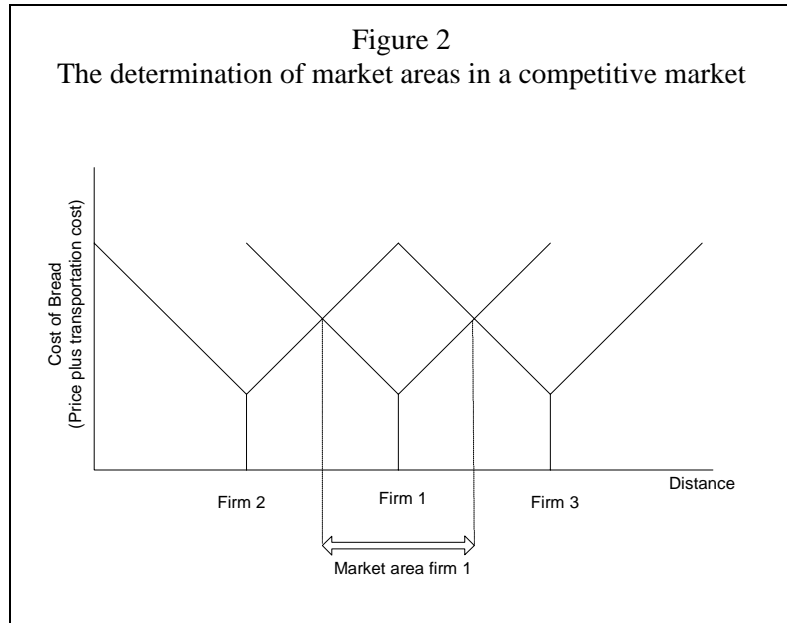
Imagine a region in which households are uniformly dispersed across space. How would retail stores, say bakeries, locate throughout this region in order to satisfy the demand for bread and baked goods in the region? If there is no bakery in this region at all, an entrepreneurial baker would locate in the most central part of the region in order to attract customers from all over the region for whom the benefit of purchasing fresh bread at the store exceeds the total cost of the purchase, which includes the monetary cost of purchasing the bread plus the cost of traveling up and down to the bakery. Of course, since the baker is the only supplier in the region, he can set his prices above competitive market prices because he is essentially a monopolist. However, if the region is

⁵ Location theory is the field of economic study that considers the structure of markets across geographical space. Location theory is based on a theoretical framework in regional economics which were developed by August Losch (1954) based on early work by Walter Christaller.

sufficiently large, there will be large areas in the region in which households have a demand for bread, but whom are unwilling to pay the high prices for bread charged by the baker, or for which the travel cost to the bakery is prohibitive. In a competitive market economy, bakeries would continue to enter the region and locate across the region's territory in order to maximize profit, spreading out across space in such a way that would attract the greatest number of customers given the location of the other providers.

Of course, as firms enter the market, households will decide to buy their bread from the bakery for which total cost (monetary plus travel costs) is minimized. Since all bakeries face a similar cost structure and (in the absence of scale economies) offer the same price, households would shop at the bakery closest to them. As a result of the entry of new bakeries into the region the average service area of each bakery would decrease, and each bakery would

end up catering to fewer households thereby reducing profitability. In equilibrium, no new firm would be able to find a market location that would be profitable and entry of new bakeries into the region would stop. A firm's market area or service area would then be defined as the geographical area over which a firm can out-price its competitors based on its net price (Figure 2).



Although the example above relies on some restrictive presumptions, location theory nonetheless provides a relatively realistic picture of the process of firm location and the resulting structure of firms across space. Based on this theoretical framework, the size of the market area for a good depends on:

- Level of travel costs: If travel costs (per kilometer or per mile) are high, this would limit accessibility to goods and services for households, thus assuring higher profit margins for new entrants that can locate closer to these households. Higher travel costs will thus result in more stores and smaller market areas. Travel costs are higher (and thus the need for firms to locate closer to household are

- higher) for goods and services that are purchased more frequently in a given time period, such as fresh dairy products, produce or meat.
- Per capita demand: If households place a high demand on a product (versus products for which demand is lower), this will increase the density of demand for the good, which in turn will result in higher profitability of each store. Higher profitability will encourage more entrants into the region and reduce market areas. Therefore, higher household demand will result in more stores and smaller market areas.
 - Population density: Similarly, higher population densities will increase the density of demand for goods, again resulting in higher profitability of each store. As a result, higher levels of population density will result in more stores and geographically smaller market areas.
 - Economies of scale. The analysis becomes slightly more complex if the production or provision of a good is subject to scale economies,⁶ although it is clear that the larger the scale economies in the production or provision of a good, the fewer firms will enter the region and the larger the market area would be. In the presence of scale economies, firms with larger market areas are able to offer lower prices at the store by taking advantage of scale economies. In fact, households might even by-pass smaller firms that are closer and incur higher travel costs in order to buy their purchases at lower prices at a larger retailer. Thus in the presence of scale economies, smaller firms would be at a cost disadvantage. As such, the larger the scale economies in the production of a good, the fewer firms would enter a region, since households make the trade-off between incurring extra travel costs to shop at a larger and cheaper outlet rather than purchasing a product closer to home at a smaller store but paying a higher price at the store.
 - Economies of scope. If households are able to purchase several different goods and services at a single location, they will be able to reduce their overall travel costs. This will encourage larger stores that provide multiple goods and services servicing large market areas.

Implications of location theory for subnational government structure

The above example of location theory implies that there are a variety of forces at work that influence a market structure and the size of market provision areas. At this point, drawing an analogy between the market example and the provision of local government goods might be appropriate, especially since many goods and services provided by subnational governments actually are private goods. Consider that instead of bakeries, the example had dealt with public school location. In fact, we could further abstract away

⁶ Economies of scale or scale economies are terms that indicate that as the scale of production increases, the cost per unit of product decreases. In other words, the existence of scale economies in production implies that smaller firms face higher costs per unit and therefore would have to charge higher prices.

from local government finances for the time being and presume that all households pay school fees directly to each public school to cover the cost of local schooling.

We should recognize that the mechanism of entry into the region, in general, would not be the same for bakeries and schools, as public schools do not enter a market or region based on the profit-motive. However, social planners (i.e., policy makers and government officials) would have to recognize the same basic trade-offs that exist in the market arena: accessibility is increased by increasing the number of schools in a region, although smaller schools are often not able to capture the same scale economies and offer the same quality as larger schools that serve a larger population and area.

Location theory also provides initial insights into the structure of local governments. Based on the principle that it is preferable to establish governments at the lowest possible level that is efficient (a principle which will be discussed in greater detail below), one could use the “provision area” (analogous to the “market area”) as the fundamental unit of government at the lowest level of government possible. If taken at face value, this would mean that the provision area around each school or public service facility could be turned into a local government with elected representation in order to assure that each local government service is provided and financed.

3. Relating the conceptual framework to subnational government structures and jurisdiction size in the real world

For policy purposes, concepts and theories are only of limited value if they are not able to link back to the reality in which government policies have to be formed. What concrete guidance does the theoretical framework provide for the policy discussion on the optimal size of jurisdictions and optimal subnational government structure? This section considers to what extent the conceptual framework has been validated and supported by empirical research, considers limitations of the conceptual framework, and explores extensions of the framework to better link it to policy practice.

3.1 Considering the validity of the overall conceptual framework

In the foregoing discussion, scale economies have been assigned an important role in the determination of an optimal structure of local government. In fact, a series of local government boundary reforms across much of Europe since WWII was driven by the idea that local government efficiency would be enhanced by the creation of larger local authorities (Martins 1995). Thus, efficiency was typically understood in the sense that the same level of public good is produced with lower costs. However, from an empirical point of view it has proven difficult to define the cost structure of locally provided public services, as it is very difficult to account for qualitative aspects and externalities in public

service provision. These studies often fail to produce a clear-cut answer regarding the optimal size of local jurisdictions.

Nevertheless studies have been undertaken in many countries to determine the minimum efficient scale of local government services by measuring only direct costs and assuming that the quality of service provision is invariant. The optimal size for local government jurisdictions seems to differ by type of service and varies across countries as a result of other conditions, including differences in geography and labor market conditions. The reliance on multi-purpose local governments in many countries further requires assigning weights to different local government functions in order to arrive at a single minimum efficient scale. The best that such studies could do was to help determine local authorities' minimum viable size threshold for a given basket of services. Depending on the type and functions of local governments in a particular country, this threshold varies from 2,000 inhabitants in Bavaria, Germany to 5,000-6,000 inhabitants in Denmark to 6,000-7,000 in the Netherlands, to 8,000 and 15,000 inhabitants of rural and urban areas respectively in the Saarland, Germany (Martins 1995).

Although the empirical literature on the provision of local government goods is too extensive to be adequately summarized here, there seems to be a consensus among economists on a number of issues. The general conclusions of the literature appear to be that:

- Scale economies exist in the provision of local government goods and services, but the level of scale economies differs greatly for different kinds of locally provided goods, and the level of scale economies can differ greatly across and even within countries. Nonetheless, typically it is possible to determine a (rough) minimum size threshold for a local government that is needed for a government unit to be able to produce a certain good or service efficiently. International experiences with jurisdiction size and minimum efficient scale are discussed in the next subsection.
- When several local government jurisdictions are available, households display a tendency to sort themselves with regard to income and/or the demand for local government services. This is consistent with the Tiebout hypothesis. However, even when large numbers of competing jurisdictions are available, large within-jurisdiction variations in income and preferences continue to persist over time, thereby weakening the implications of the club goods model.
- As the Tiebout hypothesis suggest, it appears that local government services (such as local public schools) feel some competitive pressure to provide services in an efficient manner. However, elections and “households voting with their feet” (by moving out of inefficient jurisdictions) are slow and imperfect mechanisms, allowing considerable inefficiencies to occur at the local level, even when multiple alternative jurisdictions are available.

In other words, the empirical literature regarding the optimal size of jurisdictions and the optimal structure of local government is predominantly consistent with the conceptual framework and thus supportive of its implications. However, in reality many of the assumptions upon which the conceptual framework is built do not hold strictly, and therefore we should expect the local government structure to only loosely reflect the implications of the conceptual framework, while particular characteristics of countries or regions may all effect the actual local government structure. Thus, while economic theoretical considerations provide some guidance in framing the overall local government structure, the optimal structure of local governments can vary from region to region and country to country, and will depend on the kinds of goods and services provided by local governments, the cost structure of the local government goods and services, the political and institutional environment, as well the demographic and geographical conditions.

3.2 Size of subnational jurisdictions: International experience

Several trends can be observed in different parts of the world with regard to the size of local governments. Much of the available literature in this field considers local government structure in developed economies.

During the early 1950s and up until the mid-1970s, many Northern European countries undertook consolidation of local jurisdictions in order to achieve a minimal size of basic authorities between 8-10 thousand inhabitants (Norton, 1995, p. 38). This reorganization was prompted with the perception that new technology of public services required higher levels of public investment. In addition, a meaningful size of local authorities was seen as a condition of decentralizing the power from the center. However, after the consolidation took place, advisory and executive authorities were created at sub-municipal level to bring public administration closer to the citizenry.

On the contrary, in Southern Europe, local opposition to consolidation, sometimes supported by court rulings, preserved small authorities. In those countries, (small) local governments had to achieve technical efficiency by means of inter-municipal cooperation, upward delegation of functions, and contracting services out. It is suggested that Southern European countries (France, Greece, Italy, Spain) are less prone to restructuring because they have a fused hierarchy system of central-local relations, in which delivery of local services is often ensured by central government field offices (Martins, 1995). Among transitional countries, Lithuania undertook reforms that abolished the lowest level of the Soviet hierarchy and established municipalities at the rayon (district or county) level in 1995.

As of late 1980s, many developed countries vary widely in the sizes of their basic authorities. France has the smallest average size of basic authority (1,560 persons) and Great Britain has the largest average size (122,000). Germany and the United States stand between Southern and Northern Europe in the size of their local authority (7-9,000). In those European countries that have large basic authorities (e.g., Sweden, Denmark, and

Netherlands) sub-municipal area councils provide advice on local needs, planning issues and in some cases to undertake minor responsibilities. Similarly, in the United States and Britain local administrative offices were created in a number of cities to establish closer consultative arrangements with local residents. Transitional countries also show significant variation in the size of the average population of the lowest tier of local government. The smallest size is observed in the Czech Republic (1,659 persons) and Hungary (3,241 persons). At the other extreme, the average size of basic authority in Bulgaria is 31.6 thousand persons, while in Lithuania it is 66.3 thousand persons. Obviously, these averages hide tremendous variations in the scale of basic authorities within each country. Thus, within all countries with the size of basic authority below 10 thousand, there is a large gap between a few dozen large municipalities (more than 100,000 inhabitants) and a vast majority of very small entities (fewer than 5,000 inhabitants).

The experiences in developing countries with respect to local government structure and jurisdiction size have been subject to less systematic attention in the empirical literature. Unlike many developed countries where the municipal government level tends to be dominant in providing local services, developing countries tend to have two important local government levels: one at the district level and one at the village (or sub-district) level. While the division of responsibility between these two local government levels varies from country to country, an overview of selected countries gives some idea about relative jurisdiction sizes in developing economies. For the provision of local government services, Tanzania relies predominantly on the district government level (with an average population size of 300,000) for the delivery of local government services, while contemplating a stronger role for the village level (average village population: 3,500). Districts in Malawi play a similar role, and have an average population of over 400,000. Likewise, Bangladesh and Indonesia rely almost exclusively on higher-level local governments (the Upazila level in Bangladesh, average population: 300,000; the Regency level in Indonesia: average size: 670,000) while relegating only a minor role for lower-level governments, including Union Councils (average size: 30,000) and villages (average size: 2000) in Bangladesh, and the village (Desa) level in Indonesia (average population: 3,500). On the other hand, Nepal is divided into 75 districts (average population: 360,000) and almost 4,000 villages (average size: 6,800), but the Local Government Act relies heavily on the village level for the delivery of key local services. Finally, Swaziland has a unique local government structure in that only urban areas are covered by local government jurisdictions, ranging in population from 4,800 to 67,000.

The experience in developing countries that the lowest level of local government tends to be smaller than in developed economies can be explained by a number of factors identified in the conceptual framework. Consistent with location theory, the general lack of access to automobiles in developing countries causes local residents to incur relatively higher transportation costs (time costs) to access local government services, which in turn results in smaller local government jurisdictions. However, this does not diminish the potential impact of scale economies (and the limited administrative capacity of village

governments), which appears to be a driving factor in the important role of district-level governments in developing countries.

A final observation with regard to international practices pertaining to jurisdiction and local government structure is the different trends that developed and developing countries seemed to have followed in recent years. In general, developed economies seem to be oriented over the past decades towards ongoing consolidation of local government jurisdictions. In contrast, many developing and transition countries that have been engaged in decentralization reforms –particularly in the early 1990s- seem to have tended towards greater fragmentation of their local government structure and smaller local government jurisdictions. Examples of this trend have included Russia, Hungary, Nepal, Uganda, and South Africa.

3.3 Limitations of the location analogy

While we believe that the theoretical framework –in particular location theory- provides useful insights into the optimal structure of subnational governments and the optimal size of jurisdictions, the discussion so far has presented a significant simplification of reality and has ignored several real-world complications. Three important limitations of this conceptual framework have not been addressed, notably:

First, location theory considers the cost of providing a good at certain locations, but ignores possible scale economies that might be captured by centralizing certain phases of production. For instance, to continue the bakery example, a network of associated bakers that jointly purchase intermediate inputs and business services might be more efficient in producing bread more cheaply than a network of independent full-service bakeries, where the entire production process is done in-store. This analogy extends into the provision of local government goods and services: there may be limited scale economies to the practical operation of an individual school, but there may be significant savings due to scale economies from managing and operating a number of schools that may not be captured by an individual school. For instance, it could arguably be considered duplicative, costly and inefficient to have each individual school develop its own curriculum, have its own hiring policies, operate its own payroll, have its own legal representation, administrative apparatus, and to bear the cost of local government elections for each individual school provision area. If there is truth to these arguments, the school “provision area” around a single school might not be the lowest level at which public education could be provided efficiently. As a result, local governments would have to be established to encompass several schools in order to capture these additional scale economies.

Second, information cost and transaction costs play an important role in determining the optimal structure of local government services. The larger the local service area becomes, the harder it is for local residents to have access to local government services, but also the

harder it becomes for local residents to hold their local government officials accountable.⁷ As local governments and their constituent bases become larger, there is a risk that local government officials are no longer held accountable by their constituents. This would reduce the efficiency with which local government services would be delivered. Similarly, the larger a local government's service area, the more costly (and less efficient) it would be to provide adequate information about the local government's operations to all voters/stakeholders in an area. The impact of this factor varies greatly across countries with the level of development, with local government authorities in lesser developed economies generally having a much harder time to exchange information and communicate with local residents.

Third, the basic conceptual environment presented here suggests that there exists a separate market structure for each separate good. Of course, this is not the case in reality. Similarly, scope economies often exist in the provision of local government goods and services: it is cheaper and more transparent for one general-purpose local government to administer the provision of a range of goods and services (such as education, primary health care, library services, garbage removal, public parks and amenities, and so on) than for specific-purpose local governments to be established for each individual government service.⁸ But note that while scope economies reduce the cost of providing local government goods and services, the establishment of general-purpose governments does reduce the link between the provision areas of publicly provided goods and their beneficiaries, because the service areas of optimal service areas of different locally-provided goods may vary between the different types of government services.

Finally, the framework that has been outlined presumes that households are distributed across space already, and that the location of firms (or the provision of government services) has no impact on location decisions of households. This is clearly not the case in reality both households and firms are likely to be mobile. As such, the "real world" is much more complex than can be captured by a simple representation of location theory, so we need to recognize the existence of such trade-offs and interdependencies.

⁷ Panizza (1999) adjusts the firm location theory to the public sector setup. Besides, mere travel costs, the distance to the jurisdiction center is also allowed to capture the average proximity in taste to that of the median voter. Rather than profitability or a market share in the private setup, it is the size of the public sector that drives the government to seek taxpayers' satisfaction with public goods provided. Thus, a revenue-maximizing central government faces a trade-off between its share in the public sector and the total size of the public sector. A gain in the total size of the public sector is theorized to result from moving decision-making on the type of public goods closer to the tax-payers, thus making them demand more of the public good whose type is closer to their preferences.

⁸ Transparency is relevant in that there is an information cost for households to having more government levels. In the case of multiple local governments, households would have to spend more time being informed in order to hold governments accountable.

3.4 Democratic governance and local government structure

In addition to the general limitations of local theory, an additional influence on the structure of subnational government that has not been adequately addressed so far is the political desire to have participatory and democratic governance processes. While often overlooked by economists, policy experts need to take into account this political dimension of local government structure, especially to the extent that the presence of elected bodies at the subnational level should increase the level of monitoring and accountability of subnational expenditures. As such, historical precedent and institutions cannot be ignored altogether in the formation of subnational government structures, but neither can the perception that smaller government jurisdictions are inherently closed to the people and thus more democratic.

However, participation and accountability also have a direct impact on optimal jurisdiction size. For instance, in the context of Buchanan's theory of clubs, if greater club size is associated with declines in accountability, this would result in a smaller optimal size for jurisdictions. This can be visually verified by considering Figure 1. Less accountable local government would result in a sharper decline in the benefits received by the club's members as the club size increases than would otherwise be the case; the steeper Total Benefits curve in turn reduces the optimal jurisdiction size n^* . Panizza (1999) finds a similar result by introducing accountability into the location theory model.

This link between jurisdiction size, accountability and participation is not only theoretical. For example, based on data from Denmark (both at the municipality and individual levels), Mouritzen (1989) examines the impact of size on attitudes towards local government. He finds that citizens in small jurisdictions hold more favorable attitudes toward participation and democracy, and he further finds that small local government units are more homogeneous and in fact more efficient in the provision of services.

3.5 Some more thought on clubs with geography

The conceptual theories surveyed in the previous section identify separate trade-offs that have to be considered in the design of an optimal structure of local governments: transport costs versus economies of scale, economies of scale versus congestion, allocative efficiency versus economies of scale. For each of these trade-offs, some optimal arrangement is expected to exist that would balance out the competing criteria. Moreover, both Tiebout and Buchanan identify conditions under which market forces (user charges and competition among localities) can lead us to this optimal arrangement.

Unfortunately, attempts to synthesize these separate theories run into certain constraints. First, introducing transportation costs to profit-maximizing clubs would make the revenues generated by user charges at the optimal level fall short of the provision costs by the amount of land rent from the proximity to the facility; this is known as the Henry

Jorge rule.⁹ In the case of a single public good, the optimal allocation can be achieved through market forces if the clubs own the land where their patrons live and maximize the total of user fees and land rent. However, in the case of several overlapping clubs, the optimal allocation cannot be achieved through market forces, as rent-sharing among several clubs would create incentives for each of them to under-provide. This inefficiency can be resolved by either combining all overlapping clubs into a single authority (e.g., a metropolitan government) or allowing the central government to redistribute land rent among the overlapping clubs.

The appeal of one solution over another depends on the relative advantage of the market forces over the higher-level legislature in arbitrating interest of individual sub-areas. If the market forces are assumed to be more efficient than the local jurisdiction should be established so that it would completely cover the benefit area. Any “externalities” created by the jurisdiction territory exceeding the benefit area can be “externalized” through the taxation of land rent in the benefit and redistribution towards areas that do not benefit from the public good but bear the costs through non-benefit taxation (Olson, 1969). If, however, the interests of different sub-area can be more effectively arbitrated in the higher-level legislature (e.g., matching grants), then the jurisdiction territory should be designed so that it would not exceed the benefit area and any externalities created would be compensated through intergovernmental transfers.

4. Determining the optimal number of subnational government levels

If you were to narrowly apply the economic theories guiding the optimal size of subnational government units, this would result in a system of subnational government where each subnational good or service is provided by special-purpose districts created exclusively to match the optimal service-delivery area for each government service. Yet, a number of factors prevent such a patchwork of special-purpose subnational government jurisdictions from arising.

4.1 Economic arguments in determining the optimal number of subnational government levels

The conceptual framework outlined in the previous section, including location theory, subsidiarity, the theory of clubs, and the Tiebout hypothesis, suggests that there are trade-offs in determining the optimal structure of local governments, between positive and negative features of adding an additional level of government. The primary trade-off suggests that, on one hand, a structure of subnational government that includes a local government level with small local government jurisdictions may fail to capture

⁹ See Hochman et al. (1995) for a further discussion of this issue.

economies of scale in the provision of public goods and services, while on the other hand, relying on larger jurisdictions would potentially reduce the level of allocative efficiency (by reducing the match between households' preferences and the level of service provided) and reduce the competitiveness between local governments. However, the trade-off between technical and allocative efficiency is only one of several economic factors that determine the optimal number of subnational government levels. The economic arguments in favor and against adding an additional level of government in a government structure (or further fragmenting the lowest level of government) are contained in Table 1.

Although inclusion of a lower level of government may increase allocative efficiency and bring government closer to the people, relying on an excessive number subnational government levels or even excessive reliance on special-purpose local governments does come at a cost. For instance, if two local government services could be more cheaply administered by one government level instead of two, having different government levels to provide different services may prevent the subnational government structure from capturing such scope economies in the provision of different government services. Similarly, an excessively large number of government levels often causing duplicative administrative effort at different levels of the government hierarchy.¹⁰

Table 1 Potential positive and negative effects of adding an additional level of local government	
Potential Positives	Potential Negatives
<ul style="list-style-type: none"> ▪ Achieves more homogeneous jurisdictions at the lower level ▪ Better match in provision areas for goods with small benefit areas: improves allocative efficiency ▪ More competition between local govts (Tiebout): increases in administrative efficiency ▪ Government closer to the people: increased political participation and a better informed consumer of local public goods 	<ul style="list-style-type: none"> ▪ May be technically inefficient by failing to attain scale economies in provision and administration of services not feasible at the lowest level ▪ May decrease scope economies in administering public goods: technical inefficiency ▪ A greater number of government levels will cause vertical administrative duplication and a greater probability of leakages between the vertical links in the system ▪ Excessive numbers of government levels may increase information costs for

¹⁰ Public expenditure tracking surveys have been used to show that the more links in the chain of the government's administrative structure, the greater the administrative duplication, leakages and inefficiencies that are caused (Dehn, Reinikka, and Svensson 2002).

	citizens and voters, reduce participation, transparency, and accountability
--	---

Furthermore, the information cost required for local residents to hold the various government units accountable favors a more harmonized government structure; a greater number of government levels increases the information cost for households and may in fact decrease accountability, as households are forced to spend more time being informed in order to their hold governments accountable.

While the economic theories that frame the question of optimal government structure provide guidance on the factors that determine an optimal subnational government structure, the current economics literature do not provide us with a single, comprehensive model that is able to authoritatively answer how many levels of government should exist in a country, or at which level (e.g., region, district, town, or village) subnational governments should be established. Besides, in addition to strictly economic arguments addressing the efficient delivery of services, the determination of subnational government structure is impacted by the role of regional governments are an intermediary government level, as well as by political factors, as discussed below.

4.2 Administrative constraints and the role of regional governments as an intermediary government level

Separate from the scope economies issue, there is perhaps an even more fundamental administrative dimension in determining the optimal number of government levels: one should recognize that each government level is limited in its capacity or ability to administer or deal with jurisdictions at the government level below it. Thus, in addition to being tasked with the delivery of regional government services, regional governments (as well as “higher-level local governments” such as counties, local government districts, or sub-regions) may also function as an intermediary between central and local governments. In fact, in some systems the predominant role of regional governments is to serve as an intermediary government level (e.g., provinces in the Netherlands or Zilas in Bangladesh).¹¹

As a result, the answer to the question how many subnational levels of government are needed is determined by numerous factors, including the size of a country (both in terms of population and land area), the administrative capabilities of the center and the relative development of a country’s transportation and communications networks. For instance, if

¹¹ In other systems, the intermediate government level specifically has little or no oversight over local government units within their jurisdiction. This is well-illustrated by South Africa, where the constitution speaks of national, provincial and local spheres (i.e., not levels) of government, which are considered distinctive, interdependent and interrelated, but not (necessarily) hierarchical.

the travel distance between central government and local governments is excessive for local government officials to receive direct instruction from the center, then an intermediary level of government may be warranted. Likewise, depending on the system of intergovernmental fiscal relations, the central government may simply not have the administrative capability to deal with many –say, hundreds or even thousands- local governments units.

Based on this argument, in less developed governance structures, where intergovernmental relationships are typically politically negotiated and more discretionary rather than formula-driven and institutional, there is a limit to how many lower-level government units the central government can deal with. As a result, a larger number of hierarchical tiers of government are needed to distribute the central government's reach across the national territory. Similarly, in larger countries (geographically and population-wise), all else equal, more tiers or levels of government are generally needed. As governance techniques become more developed with the introduction of information and communication technologies, fewer tiers would be necessary and larger government units can be governed.

Intermediate levels of government can either be established as deconcentrated tiers of the central government, or as separate regional government levels with elected officials. The choice whether to establish an additional government level often depends on politics as well as the exact functions assigned to intermediary government levels, including whether this government level has supervisory responsibilities for lower-level local governments. Obviously, the greater the responsibilities of a level of government, the greater the need for direct, downward accountability to the voters.

4.3 Political considerations

There might also be political, socio-economic and even cultural factors that promote having intermediate levels of government. In some countries, the establishment of a regional government level is shunned –or the responsibilities assigned to this government level minimized- in order to prevent a vertical power struggle between the regional governments and the center. Likewise, regional governments may be overly fragmented in order to reduce the political power of regional executives. Countries where vertical power considerations seem to have played a crucial role in the subnational government structure include Nigeria (states appear to have been fragmented for political purposes); Russia (see Martinez-Vazquez and Boex (2001) for a detailed discussion of vertical power issues in the formation and evolution of Russia's subnational government structure. More recent political reforms were introduced to further curtail the power of regional governors (Mosnews, 2004)); and Nepal (districts were excessively fragmented, and expenditure responsibilities largely circumvented the district level) (Arze, 2004).

Certain regions might be distinct, and it may be politically more expedient to deal asymmetrically with different parts of a country by inserting an intermediary level of

government (Bird, 2003). In fact, the introduction of regional governments in Wales and Scotland in the United Kingdom, as well as the introduction of a regional authority in Corsica in France, and the special autonomy of Aceh and Irian Jaya (West Papua) in Indonesia speak to this point. Asymmetry in intergovernmental relations is discussed in some greater detail in Section 5 below.

5. Horizontal and vertical relationships in the subnational government structure

Once the framework for a country's subnational government structure is defined by the number of government levels or tiers and the size range of subnational jurisdictions at each level of government, the remaining dimension of the subnational government structure is the definition of the vertical and horizontal relationships between levels of government and subnational government units, respectively. As such, this set of policy questions considers issues such as how subnational government jurisdictions are formed and governed and how subnational governments relate to the central government and each other. By necessity, this third dimension of subnational government structure is more political in nature, and varies widely from country to country based on historical and institutional differences.

In our attempt to cover the main issues related to the horizontal and vertical relationships in subnational government structures, this section will address four specific policy questions. First, we address the fact that the philosophical underpinning of a decentralized system has a crucial impact on the subnational government structure and on the relationship between different government levels. Second, there is the question of subordination of subnational governments: to what extent are local governments in fact treated as political agents of the national government, as opposed to autonomous subnational economic clubs that are allowed to arrange their own affairs? Third, how are the boundaries of subnational governments defined? Fourth, to what extent can symmetry or asymmetry enhance the system of intergovernmental relations?

5.1 Philosophy underlying subnational government structure

Many of the questions that relate to the structure of subnational government start with the basic question how subnational governments should be formed: should the subnational government structure be structured in a bottom-up process through the formation of local communities and local charters of self-governance, or should subnational government structure be imposed top-down from the center? International experience presents two general approaches to local government: (1) community-based (bottom-up) local government formation and (2) utilitarian (top-down) local government administration.

Community-based local government is understood as government formed by and for citizens of the community, following the belief that people who belong to an area have the right to participate in governing themselves as a collective (Norton 1995). Home rule thus not only allows communities to decide which local government services to provide at the local level, but it allows the local community to define the size and composition of the “club” that it may wish to organize. This approach is especially popular in Denmark and the Nordic countries. Local self-government or home-rule is also the norm in many U.S. states for municipal governments (Katz 1999).¹² Similarly, community-based local government is common throughout francophone West Africa.

By contrast, “utilitarian administration” refers to a government structure that is imposed top-down in order to assure stability and to create an enabling environment where citizens have opportunities for self-realization. This approach is prevalent in Anglophone countries and many countries of the British Commonwealth; the primary motive for the structure of local British colonial administrations was to maintain peace and order at the local government level. Once established, these local governments were primarily responsible for their own development.¹³ The top-down administrative structure also allows the center access to well-structured local counterparts for the implementation of top-down national programs or policies, when necessary.¹⁴

It can be argued that a utilitarian administrative structure complements a market-based “laissez-faire” state, where the top-down establishment of local government jurisdictions is justified in terms of universal coverage and economic efficiency, as long as the operation of local government affairs is free from arbitrary central interference. As such, this approach is the closest to the club theory model where patrons take the set of available clubs as given, but can freely choose which to join. Of course, more authoritarian governance systems (including socialist planned economies) also typically rely on a top-down approach in defining the subnational government structure, but under those circumstances, subnational jurisdictions often have little discretion over their activities, which are typically controlled by central authorities in a hierarchical manner.

In Central and Eastern European countries, although there were some systems of local government before the transition, all subnational power structures had been mostly authoritarian and provided little opportunities for self-government. Therefore, the process of transition and democratic institution-building required the creation of a new system of

¹² Although home-rule local government systems are common for establishing municipal governments in many U.S. states, the approach typically coexist with a more top-down system of county government.

¹³ For instance, the Union Parishad structure and the legislated role of the Union Councils in Bangladesh essentially follow key elements of the original local government structure imposed by the British.

¹⁴ In developed countries and post-colonial developing countries, the expenditure responsibilities and revenue assignments of many local governments were modified over time. Although initially most local governments functioned as autonomous clubs the financial responsibility for key social services such as primary education and other basic services was typically centralized, along with key revenue instruments. Subsequently, many countries delegated the responsibility of service provision in these areas back to the local level, funded by conditional grants.

local government. Most transitional countries in Central and Eastern European countries adopted the Council of Europe's approach to local governance and signed the European Charter of Local Self-Government.

A further tendency that can be observed internationally is that countries with a utilitarian approach to government tend to favor larger-scale local authorities with fewer –if any– intermediary government tiers. As noted above in Section 3.2, many developing countries that were influenced by British colonial structures have a utilitarian administrative structure based on utilitarian administrative structures at the district and village level. At the same time, countries with community-based local government tend to have more complex multi-tier systems with large number of small basic authorities (Norton, 1995).

5.2 Subordination of Local Governments

To the extent that the responsibilities for the delivery of government services are not assigned to the local government level, it may be necessary for the central government to organize itself in a deconcentrated fashion in order to deliver the public services for which it remains responsible. Similarly, centrally-appointed local officials or commissioners may be posted at the regional or local level to represent the center and to monitor local affairs. International practice offers several possible arrangements for the intermediate local authority (Norton, 1995):

- A state-appointed official responsible for state functions over a large area without an elected council of its own (e.g., Bulgaria and Lithuania).
- A state-appointed official for each local jurisdiction who acts as executive and presides over an elected council (e.g., centrally-appointed Mayors in the Netherlands).
- A prefect who is responsible for state functions in the intermediate local jurisdiction in parallel with an autonomous assembly which provides major services and elects its own chairman and executive.

In Council of Europe countries, written constitutions guarantee a system of local governance and give it a secure position within the national system. In the United Kingdom, although the rights of communities to self-government received general recognition, the 'unwritten constitution' affords no protection to them. This became clear under the Thatcher government that in various ways decreased discretion of local authorities and prevented them from increasing the quantity and quality of their services. Similarly, federal constitutions of the United States and countries in the British Commonwealth (including Australia and Canada) make no specific provisions protecting local governments. However, constitutions of constituent states typically do recognize local authorities and protect them from arbitrary dismissal or state interference. In continental Europe, constitutions of member states of federations and regional governments often detail competencies of local governments, as in Switzerland,

Germany, Italy, and Spain. In Japan, the 1947 Constitution recognizes local government as a part of the new democratic administrative system based on the rule of law.

In most developed economies with established democracies, multi-tier systems of local government are generally not hierarchical in the sense that there is no political control of one tier over another. Also, in some transitional countries (e.g., Hungary, Poland, and the Czech Republic) tiers of local government are non-hierarchical. In contrast, the former Soviet system was completely hierarchical, where lower-level local governments were considered an integral part of the higher-level government jurisdictions, to the point that the budgets of lower government levels were incorporated into the budget of the next-higher government level. This has been referred to as the “matrushka” approach to subnational government structure, based on the name for Russian nesting dolls, where each doll neatly fits within a slightly larger doll.

While in many of the developed countries it is now a principle that all tiers of local government enjoy a fairly high degree of autonomy, relationships between different levels of government are still quite close in many unitary countries in continental Europe (Norton 1995). This has been the result of persuasion, incentives, and sometimes the result of legal requirement. Thus, the *Local Government Law* in Italy repeatedly requires inter-level consultation. In France, communes were obliged -often by fiscal pressure from the State- to join together in a single- or multi-purpose syndicates sometimes vested with taxation powers. In Sweden local government legislation requires sharing responsibility between authorities for regional and local transportation usually carried out through jointly-owned companies. In Finland, joint municipal boards are established with responsibilities for hospitals, care of the disabled, and regional planning. In the Nordic countries local authorities readily cooperate with one another and with other institutions to provide services for their residents (Hughes, 1998). In each of the fifty-eight state administrative units in Slovenia, municipal councils appoint their representatives to the advisory committee that ensures coordination between municipal bodies and state administrations.

The European Charter of Local Self-Governments requires that any administrative supervision of local authorities may only be exercised in cases determined by the statute and aim only at ensuring compliance with the law or with regard to expediency in respect to delegated tasks. Administrative supervision is reduced to a simple supervision of legality in Spain, France, Portugal and Sweden. In principle this is also the case in Denmark, Finland, Norway and Poland, but exceptions are tolerated: in particular sectors or under special laws (Denmark, Finland) and in matters which are the subject of “delegated powers”, as in Poland (Council of Europe, 1994). When supervision of expediency is envisaged, considerations of “economy” (Poland, Austria), “efficiency,” “adaption” or “utility” (Austria) or “strictly necessary situations” (Netherlands) could intervene.

The record in developing economies on vertical subordination is mixed, but generally elected local governments are generally much weaker with greater state control over local affairs. For instance, it is not unusual at all in developing countries for the executive of district-level governments to be appointed by the center, despite the presence of well-functioning elected district councils. In addition –regardless whether or not district executives are centrally appointed- it is also quite common for there to exist an entire parallel central government administrative structure, sometimes all the way down to the village level). These Regional, District and Village Commissioners are often presidential or ministerial political appointees which are exclusively accountable to the center, whose sole responsibility is typically to monitor local affairs and exerts state power as directed.

While in general it is fair to say that local governments in developing countries have a more limited administrative capacity, fewer assigned responsibilities, and have more limited policy discretion than local governments in more developed countries, there are exceptions to this rule. For instance, with the reintroduction of democratic civilian rule in Nigeria in 1999, all three levels of government (federal, state, and local) have been governed by elected councils and elected executive officers, and each have substantial autonomy over the execution of the functional responsibilities assigned to each respective government level (Alm and Boex, 2002).

5.3 *Boundaries*

Practices and legal procedures for changing subnational boundaries vary tremendously across countries. For example, legislative provisions that stipulate prior consultation with communities effected by boundary changes exist in German, Danish, Finnish, French, Greek, and Polish legislation. These consultations usually take place in practice and resistance to attempted consolidation, in particular, is common. These consultations are often not authoritatively binding in any way, allowing the competent authority discretion to override them. Consultation may concern the municipality itself (Spain, Poland, Portugal), the citizens involved (Denmark, Italy, France), or simply an expert committee to the Prefect (Greece) or a special council to the Minister of Interior, comprising members nominated by representatives of the two municipal levels as in Denmark (Council of Europe, 1994).

Most countries that have carried out major structural reforms of their subnational structures worldwide have done so on a compulsory basis, usually after the encouragement of voluntary consolidation had failed (Martins, 1995), but different modalities are followed in different countries. In the Netherlands the Parliament has the exclusive power to restructure local government boundaries, and it has often done so (Hughes, 1998). In contrast, communes in Switzerland have an unwritten constitutional right to merge with other communes or to remain independent – a freedom which cannot be withdrawn by cantons. In New Zealand, a Local Government Commission was established to preside over reforms of local authorities' internal structures as well as their boundaries. In Slovenia, the National Assembly (parliament) legalizes establishment of

new administrative units along with the demarcation of their territories following a local referendum in the affected area. In Bulgaria, any decisions on the creation, transformation or dissolution of municipalities, including any changes in boundaries, are subject to local referendum.

5.4 Asymmetry

In the design of central-regional fiscal relations there has been recently more discussion about the pros and cons of symmetric versus asymmetric intergovernmental policies, which asks the question whether a federal (or central) government should treat all state (local) level governments the same, or whether there is room for the center to treat (a group of) selected subnational jurisdictions differently. This issue becomes especially relevant in light of the great differences in administrative capabilities between strong urban local governments and weaker rural local governments. A related question with respect to symmetry arises at the regional-local level: should a federal government require all regional governments to use the same approach in structuring regional-local relations?

With regard to the first question, whether central (federal) governments should treat all local (regional) governments the same, international practices again reveal a variety of approaches. Two policy choices can be defined in this regard. In one case, all local governments would have exactly the same expenditure responsibilities. This of course would require a relatively high degree of consistency in the size and administrative capabilities of local governments, a case likely reserved for small, homogenous unitary countries. In the other case, local governments with different characteristics would be classified according to some scheme and assigned different expenditure responsibilities and revenue-raising authority based on their size and administrative capabilities. The latter is the common approach among international practices. For instance, in Spain legal provisions assign different degrees of responsibility to local governments according to their size. Likewise, the assignment of additional expenditure responsibilities to special or autonomous jurisdictions within a country is also quite common in international practice, particularly for large metropolitan areas as well as for subnational regions with distinct ethnic, linguistic or cultural identities.

With regard to the second question (whether a federal government should require all regional governments to relate to their local government in the same manner), many federal structures have opted to allow regional governments varying degrees of discretion in determining local government structure. In the United States and Canada, local governments are creations of the state government. In Germany, *Länder* have less discretion but still a variety of local government structures exist: some regional governments have a local structure with municipalities of a small average size (e.g., Mecklenburg-West Pomerania), while others have a local government structure with relatively large local governments on average (e.g., North-Rhine Westphalia). In fact, in a number of federal countries, local government structures are mixed in a sense that the

local government structure and number of local government tiers vary between different regions of the country (for instance, different states in the U.S. have different local government structures). Nevertheless, in general the structure of local government seems to be related to the concept and role of local government in a particular country.

6. Observations from international experience

The key policy issue that this paper set out to address is to determine what guidance economic theory and principles provide in determining the optimal size of subnational governments, and the associated optimal structure of subnational governments. In way of our concluding thoughts on this topic, we would like to concretely connect these conceptual discussions to a broad overview of international experiences and trends as it pertains to subnational government structure. We do so along the line of three major observations regarding subnational government structure around the world, with a particular emphasis on developing and transition economies.

Observation 1: It is often history and politics -not economics- that determines subnational government structure

Our first observation with regard to subnational government structure is that economic theory seems to have little influence on local government structure: it is often history or political forces that shape the government structure.

In fact, decentralization reforms in many developing and transition countries –including reforms that shape subnational government structure- are typically determined either by historical governance patterns or by political factors, but not by economic considerations. While some developing countries still rely on subnational government structures similar to those inherited during colonial periods,¹⁵ other countries' subnational government structures were predominantly driven by the political desire to decentralize in reaction to previously highly centralized systems (see next observation).¹⁶ To the extent that assignments of expenditure responsibilities and revenue assignments were made to local governments that were generally too small and structurally under-capacitated to deliver local services and collect local taxes, it should not come as a surprise that many

¹⁵ It is not exclusively developing countries' subnational government structures that are determined by historical forces. The county structure of the State of Georgia has changed little for almost two centuries. The size of each Georgia county was originally determined as the areas within which a resident could travel to and from the county seat in one day by mule.

¹⁶ Likewise, the evolution of Nigeria's state structure has more to do with ethnic, religious and political balance than with economic considerations. In fact, Nigeria's transfer system seems to have been formulated to fit the objective of fragmentation rather than quality service delivery. See Alm and Boex (2002).

developing and transition countries experienced systematic problems in the early stages of their decentralization reforms.

In contrast to many developing and transition economies, it appears that economic considerations have commonly been a constructive factor in the reform of subnational government structures in many developed countries over the past decades. As noted earlier, many developed economies are engaged in a (gradual) process of consolidating local government units in order to form larger jurisdictions, which are felt to be more optimal based on economic grounds. A recent example of this trend is the ongoing reform of Denmark's administrative structure, which is heavily driven by considerations of efficiency in service delivery and the administrative capacity of smaller local governments.

Observation 2: In developing and transition countries, over-fragmentation of subnational government structure is a common occurrence

A second trend that we observe is that many developing and transition countries that went through major decentralization in the 1990s experienced an over-fragmentation of their subnational government structure early-on in the decentralization process.

Many developing and transition countries that enthusiastically pursued decentralization in the 1990s share a common thread: over-centralization during the preceding regimes and an intense desire to enhance democracy by moving government closer to the people and reducing the power of the center. Decentralization reforms that followed this pattern include Uganda's decentralization reforms after the removal of Idi Amin; Nepal's decentralization after the restoration of multi-party democracy in 1990; Indonesia's democratic "reformasi" of 1998; post-Soviet intergovernmental fiscal reforms in Russia and other transition economies; South Africa's transformation of its public sector in the post-apartheid era, and so on.

Because decentralization reforms in many developing and transition countries were based on a political reaction to over-centralization, it appears that many reforms have in fact resulted in an over-fragmentation of the subnational government structure, either by creating inefficiently small subnational jurisdictions or by assigning responsibilities to inefficiently low government levels. Once this problem was recognized in the late 1990s and early 2000s, the focus in many of the aforementioned countries began to shift to attempts to reduce the fragmentation of the lowest level of government and to realign expenditure assignments to the appropriate government levels.

Observation 3: Intergovernmental relations below the national level –including the role of intermediary government levels- tends to be under-addressed in decentralization reforms

A third and final observation with regard to subnational government structure is that relations between different subnational government levels (as opposed to central-subnational fiscal relations) tends to be a difficult and contentious issue that often has not received adequate attention in decentralization reforms around the world. Russia's decentralization reforms provide an indication of the secondary role often assigned to intergovernmental fiscal relations below the national level; while federal-oblast relations were the primary focus of intergovernmental relations throughout the 1990s, the sub-federal dimension did not emerge as a true policy issue until 2000; serious policy reforms of Russia's sub-regional local government structure did not materialize until 2004.¹⁷

The dearth of attention paid to intergovernmental relations between subnational governments is not confined to either federal countries or to transition countries; this issue is also common in many developing and unitary countries. For instance, common problems in relations between different local government levels is the definition of the respective roles and relationships between district-level local governments and sub-district level local governments, including confusing (and sometimes overlapping) expenditure assignments, problematic revenue sharing arrangements, and problems associated with the hierarchical power relations between different local government levels.

An essential challenge in government systems with multiple local government levels has been to assure accountability at the various government levels. It is not uncommon in developing countries to have directly elected lower-level local governments, but to have these local governments supervised or controlled by higher-level local governments that are either indirectly elected or not elected at all. For instance, this is the case in Bangladesh, where the elected Union Parishad is controlled by the Upazila, which is formed by a combination of central government appointees and Union Parishad Chairmen. Likewise, in Nepal the District Development Committee is comprised of chairmen of Village Development Committees within the district. The same organizational set-up is common in many other developing economies. While these arrangements originally intended to use district-level councils as a platform for coordination between lower-level governments, in many cases these subnational governance arrangements have evolved to place substantial supervisory responsibility and control in an unelected (or indirectly elected) level of government. In the end, such organizational structures tend to limit the accountability of directly elected local officials, and these structures may be exploited to result in elite capture of the local resource allocation process. The cost of this "vertical fragmentation" of power between different subnational government levels is a lower level of service delivery, administrative duplication, reduced accountability, and greater resource outflows at each link in the hierarchical chain (e.g., see Dehn, Reinikka and Svensson. 2002).

¹⁷ See the appendix for a more complete discussion of the reform of sub-federal relations in the Russian Federation.

One potential solution to the accountability problem is to more clearly define the roles of the respective government levels and to introduce directly-elected intermediary government officials, either at the regional (provincial) or at the district level. However, in many countries, a stronger role for the intermediate level seems to be resisted in part on political grounds as central authorities appear concerned over potential competition for power from strong regional or local governments. Nonetheless, the organization of relations between different local government levels, and the implications of the local government structure on local accountability and local service delivery should be subjected to substantial scrutiny as part of the analysis of any country's decentralized local governance arrangements.

References

- James Alm and Jameson Boex. 2002. *An Overview of Intergovernmental Fiscal Relations and Subnational Public Finance in Nigeria*. International Studies Program Working Paper 02-01. Andrew Young School of Policy Studies, Georgia State University.
- Arze, Javier. 2004. *Review Of The Nepal Expenditure Assignment Study*. Review for the Local Bodies Fiscal Commission. Andrew Young School of Policy Studies, Georgia State University.
- Roy Bahl. 1999. *Implementation Rules for Fiscal Decentralization*. International Studies Program Working Paper 99-1. Andrew Young School of Policy Studies, Georgia State University.
- BBC News. 2003. "France offers Corsica new deal". Friday, 4 April, 2003.
<<<http://news.bbc.co.uk/2/hi/europe/2916247.stm>>>
- Bird, Richard. 2003. *Asymmetric Fiscal Decentralization: Glue or Solvent?* International Studies Program Working Paper Number 03-09. Andrew Young School of Policy Studies, Georgia State University.
- Buchanan, James. 1965. An Economic Theory of Clubs. *Economica* 32: 1-14.
- Council of Europe. 1994. Draft Report on Monitoring the Implementation of the European Charter of Local Self-Governments.
- Danish Ministry of Interior and Health. 2004. Recommendation of the Commission on Administrative Structure, No. 1434, January 2004.
<<http://www.im.dk/publikationer/struktur_uk/index.html>
- Davey, Kenneth. 2002. "Decentralization in CEE Countries: Obstacles and Opportunities," in Gábor Péteri, ed., *Mastering Decentralization and Public Administration Reforms in Central and Eastern Europe*. LGI Books: Budapest.
- Dehn, Jan, Ritva Reinikka, and Jakob Svensson. 2002. Survey Tools for Assessing Service Delivery. Development Research Group. The World Bank.
- Hochman, Oded, David Pines and Jacques-Francois Thisse. 1995. "On the Optimal Structure of Local Governments," *American Economic Review* 85, 5: 1224-40.

- Hughes, Michael, Michael Clarke, Hubert Allen, and Declan Hall. 1998. *The Constitutional Status of Local Government in Other Countries*, Report Prepared for the Commission on Local Government and the Scottish Parliament. The Scottish Office Central Research Unit.
- Kourliandskaia, Galina, Yelena Nikolayenko, and Natalia Golovanova. 2002. "Local Government in the Russian Federation," in Victor Popa and Igor Munteanu, eds., *Developing New Rules in the Old Environment. Local Governments in Eastern Europe, Caucasus and Central Asia*. LGI Books: Budapest.
- Katz, Ellis. 1999. *Local self-government in the United States*. Issues of Democracy: An Electronic Journal of the U.S. Information Agency (Volume 4, Number 1).
- Local Government Initiative. 2000. *Decentralization: Experiments and Reforms*. LGI Books: Budapest.
- Losch, August. 1954. *The Economics of Location*, New Haven, Conn.: Yale University Press.
- Martins, M.R. 1995. "Size of Municipalities, Efficiency and Citizen Participation: A Cross-European Perspective". *Environment and Planning C: Government and Policy* 13:441-458.
- Martinez-Vazquez and Boex. 2001. *Russia's Transition Towards a New Federalism*. Washington DC: World Bank Institute Learning Resources Series.
- Martinez-Vazquez, Jorge, Andrey Timofeev and Jameson Boex. 2005. *Reforming Regional-Local Finances in Russia*. *World Bank Institute Learning Resources Series* (forthcoming).
- Mosnews. 2004. Russian Upper House Gives Putin Green Light to Appoint Governors. 8 December 2004 <<http://www.mosnews.com/news/2004/12/08/governors.shtml>>
- Mouritzen, Poul Erik. 1989. "City Size and Citizen Satisfaction: Two Competing Theories Revisited." *European Journal of Political Research* 17: 661-688.
- Norton, Alan. 1994. *International Handbook of Local and Regional Government: a Comparative Analysis of Advanced Democracies*. Aldershot, Hants, England; Brookfield, VT: Edward Elgar.
- Olson, Mancur. (1969), "The Principle of 'Fiscal Equivalence': The Division of Responsibilities among Different Levels of Government," *American Economic Review* 59(2), pp. 479-87

Wetzel, Deborah. 2001. "Decentralization in the Transition Economies: Challenges and the Road Ahead," World Bank mimeo.

Young, Stephen. 2001. Devolution in the United Kingdom: A Revolution in Online Legal Research. <<http://www.llrx.com/features/devolution.htm>>

Appendix

Reform of the structure of subnational governments in the Russian Federation

The current Russian system of government is heavily based on the hierarchical system of administrative subordination inherited from the Soviet Union.¹⁸ The government structure of the Soviet Russia had four tiers of government: (1) federal (RSFSR); (2) regional (ethnic republics, krais, okrugs, oblasts, and autonomous areas); (3) first local tier (sub-region cities and rayons); (4) second local tier (sub-city districts and sub-rayon towns, townships, and rural districts).¹⁹ During the Soviet era, each administrative unit was governed by two parallel bodies: the local committee of the Communist Party and the local Soviet (council). However, elections to the Soviets were not competitive and a single candidate for each district was effectively nominated by the Party. Thus, all decision-making was made within the Party apparatus and local Soviets were only legitimizing these decisions and implementing them through the local executive branch.

Politically autonomous local governments were introduced in the Russian Federation with the first law on local self-government in 1991.²⁰ The law introduced dual subordination of local executives both to the local council and to the higher-level executive bodies in the regions. Moreover, the law required direct election of the of local executives head. The system of local governance was further transformed with the introduction of the 1993 Constitution. The Russian Constitution guarantees local communities the right to govern local affairs separately from the sovereignty of the state. However, the Constitution established a general framework for local self-governance without any references to the existing administrative structure. The implementation of the constitutional right for local self-government is assigned to the joint responsibility of the Federal and regional governments. Thus the Constitution set the stage for the federal government to enact a number of constitutional laws which further developed the framework for local governance.

In conformity with the Constitution, the *Law on the General Principles of Local Self-Government* of 1995 allows for regional governments to determine the approach by which local self-government is instituted, meaning that regions could permit local governments to be established at the rayon level as well as at the level of sub-rayon towns or villages, irrespective of the size of their population. A complicating factor is that the law grants a uniform legal status to all local government entities. This means that all local

¹⁸ See Martinez-Vazquez, Timofeev and Boex (2005) and Kourliandskaia et al (2002) for a detailed overview of the history of Russia's local government structure.

¹⁹ Rayon is a Soviet equivalent of U.S. counties.

²⁰ Federal Law No. 1550-1 *On Local Self-Government in the RSFSR* (July 6, 1991).

governments – be they *rayons*, towns, townships of rural districts – enjoy the same institutional and administrative rights. This has been interpreted to mean that the law does not allow “subordination” of one municipality to another. Kourliandskaia et. al. (2002) thus conclude that pursuant to the *Law on the General Principles of Local Self-Government*, the entire territory of the Russian Federation should be divided into non-overlapping jurisdictions (territories) of municipalities or local self-government entities. In practice, regions had several options for superimposing the institution of self-government over the existing hierarchy of local administration.²¹

Table A			
The inherited administrative structure of territorial division			
Tier	Administrative Unit	Number of Units	Average Population
1	Central government	1	145.5 million
2	Regional level units	89	1.6 million
3	Rayon level units	2,513	57,900
4	Township level units	27,120	5,400
Note: As of January 1, 2000. Source: Goskomstat, 2001			

At present, there are about 29,500 units of local administration below the regional level in Russia (see Table A). However, as recorded by Kourliandskaia et. al. (2002) of these local administrations there are only 12,261 that are officially registered as self-government units; only 11,691 have elected representative authorities; only 11,209 municipalities are endowed with municipal property; and only 4,500 have fully independent budgets.

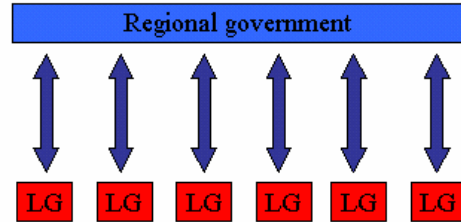
The current practices across Russia’s regions in defining local government structure does not necessarily always conform to the legislative requirements. While Russia historically had a very rigid administrative hierarchy in which administrations of smaller territories were subordinate to higher-level administrations, municipalities that have come into existence in recent years have formed at different tiers of the former administrative and territorial hierarchy. All these factors help provide a quite diverse scene.

²¹ For an overview of these issues, see Kourliandskaia et al. (2002).

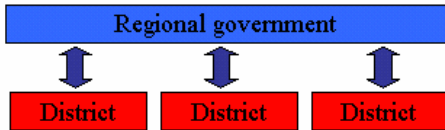
Box A
The various structures of local government in the Russian Federation

These are three examples of a **Type I** local government structure, where the regional (oblast) government deals with one level of local government, either at the district (rayon) level, at the local (town or village) government (i.e., sub-rayon) level. Note that different types of local governments should still be considered one local government level if they are not hierarchically structured.

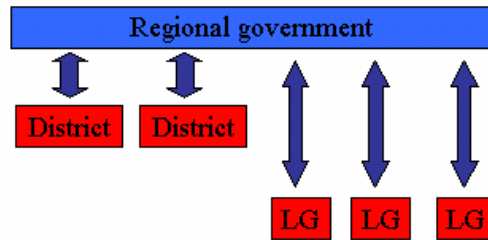
“Type I –A”



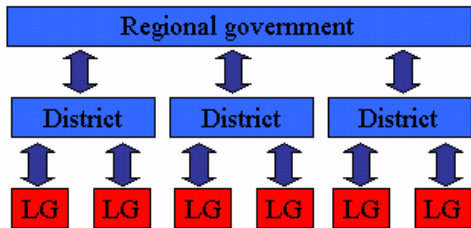
“Type I –B”



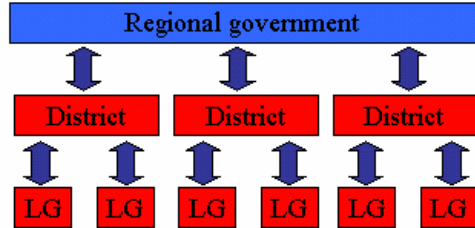
“Type I –C”



This is an example of a **Type II** local government structure, where the district level is a deconcentrated tier of the regional government and local governments are deprived of some of their budgetary rights.



This is an example of a **Type III** local government structure, where two levels of local government exist and local governments are hierarchically subordinate to the district level governments.



Box A provides examples of three basic types of local public administration currently in operation in different regions of Russia. A majority of regions established self-government at the first tier of local administration (sub-region cities and rayons). In fifteen regions self-government is established at the second local tier (towns, townships, and rural districts). In the rest of the regions self-government is established at both the levels of local administration so that some municipalities are subordinate to others.

As noted by Kourliandskaia *et. al.* (2002), the forms of local self-government organization summarized in Box A can make up different combinations in different subjects: for example, of 295 municipalities of Tyumen Oblast, 4 have been formed at the level of large cities, 2 – at the level of *rayons* with the rest set up at the level of small townships below *rayons*. In Vladimir Oblast, 18 municipalities set up at *rayon* level contain lower-level local governments while the other 7 municipal entities having a one-tier type of organization. The second-tier municipalities in the Vladimir Oblast are individual populated areas that have opted for local autonomy. However, regional authorities are unwilling to recognize them as full-fledged participants in intergovernmental fiscal relations and deal with them through the first-tier bodies of local self-government.

The most recent development in Russia's local government has been the appointment by President Putin of a high level commission (the Kozak Commission) to review and propose reforms in subnational government. The Presidential Administration introduced two laws on the reform of regional and local government respectively in early 2003, which, with strong backing from the President, were approved in September 2003. The Kozak laws essentially mandate a uniform two-tier structure of local government in the entire territory of the Russian Federation. An explicit distinction is made among rural and urban types of municipalities. At the top tier, the two types of municipalities are called municipal and urban districts, while at the bottom tier the names are rural and urban settlements respectively. This approach would be a departure from the regional variations allowed under the *Law On the Fundamental Principles of Local Self-Government in Russia*, and appears to return to the uniform, hierarchical approach under the Soviet system.

The Kozak laws also regulate the minimum and maximum size of municipalities. This is again in sharp contrast to the previous legislation, which stated that no inhabited territory could be deprived of an opportunity to exercise self-government. In the new law only those administrative units that satisfy specific criteria can become municipalities. To become a bottom level municipality, an administrative unit should either have a status of a town according to the administrative-territorial codification, or represent a rural territory with population of more than 1,000 people in one settlement, or a rural territory including several settlements regardless of their population. Territories satisfying those criteria could get the status of an urban or rural settlement accordingly. There are no

special requirements for the minimal size of population for municipal districts. However, there is an upper limit on the territory of a top-level municipality: the border should be reachable from the administrative center by transport within one day. Similarly, the center of a bottom-level municipality should be reachable by foot from anywhere on its territory within one working day.

Apart from being elected directly, the Kozak laws establish that the representative body of a municipal district can also be formed out of the council members of constituent settlements. In this case, constituent localities can complement the district own-source revenue with contributions earmarked for the provision of inter-settlement services (roads, transportation, healthcare, waste utilization, etc). The contribution of each settlement would be determined on per capita basis.

Similar to the principle maintained in the current legislation, the Kozak laws prohibit the subordination of one local government or its officials to that of another tier. However, regional authorities can participate in the selection of a municipal CEO if such a post is established by the municipality in addition to the elected municipal head. In this case the selection commission would have to be confirmed by the regional legislatures, who can nominate up to one third of the commission members.