Managing the Coordination of Service Delivery in Metropolitan Cities:

The Role of Metropolitan Governance

Enid Slack

The World Bank
Finance Economics and Urban Development
Urban Development Unit
August 2007
Abstract

This paper examines different models of governing structure found in metropolitan areas around the world. It evaluates how well these models achieve the coordination of service delivery over the entire metropolitan area as well as the extent to which they result in the equitable sharing of costs of services. Based on theory and case studies from numerous cities in developed and less developed countries, the paper concludes that there is no “one size fits all” model of metropolitan governance. Other observations from the case studies highlight the importance of the process of implementing a metropolitan structure, the need to match fiscal resources with expenditure responsibilities, the need to have a governance structure that covers the entire economic region, and the critical importance of having a strong regional structure that ensures that services are delivered in a coordinated fashion across municipal boundaries.

This paper—a product of the Urban Development Unit, Finance Economics and Urban Development Department—is part of a larger effort in the department to understand better how to improve management and financing of urban infrastructure based on country experiences. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The author may be contacted at enid.slack@utoronto.ca. or through Patricia Annez at pannez@worldbank.org.
Managing the Coordination of Service Delivery in Metropolitan Cities: The Role of Metropolitan Governance

A report prepared for the World Bank by Enid Slack
Institute on Municipal Finance and Governance Munk Centre for International Studies University of Toronto
Table of Contents

Executive Summary ......................................................................................................................................................... 2
1. Criteria to Evaluate Different Models of Government Structure ................................................................. 8
   1.1 Economies of scale ........................................................................................................................................... 9
   1.2 Externalities .................................................................................................................................................. 10
   1.3 Equity ........................................................................................................................................................ 11
   1.4 Accessibility and accountability ................................................................................................................... 11
   1.5 Local responsiveness ..................................................................................................................................... 12
   1.6 Summary of Criteria for Designing Local Government .............................................................................. 13
2. Models of Government Structure .......................................................................................................................... 14
   2.1 One-Tier Fragmented Government Model .................................................................................................... 14
   2.2 One-Tier Consolidated Government Model ................................................................................................... 15
   2.3 Two-Tier Government Model .......................................................................................................................... 18
   2.4 Voluntary Cooperation ................................................................................................................................... 20
   2.5 Special Purpose Districts ............................................................................................................................... 23
   2.6 Role of Senior Levels of Government ........................................................................................................... 24
3. Observations on Governance in Metropolitan Cities around the World .............................................................. 25
   3.1 Not only do different governance models work in different cities, models can evolve over time in any one city ............................................................................................................................ 26
   3.2 Voluntary cooperation may be necessary where the metropolitan area is large ........................................ 34
   3.3 Not all consolidated cities cover the entire metropolitan region .................................................................. 37
   3.4 Consolidation does not necessarily reduce costs ......................................................................................... 39
   3.5 The national and local context (for example, history and political culture) has a significant influence on how well a given structure will work ................................................................................. 41
   3.6 The process of implementing a metropolitan structure is crucial to the success of the outcome ................. 46
   3.7 Politics, and not economics, often dictate the ultimate structure ................................................................. 49
   3.8 Fiscal resources need to match expenditure responsibilities but rarely do ................................................... 51
4. Concluding Comments ............................................................................................................................................... 56
References ........................................................................................................................................................................ 59
Executive Summary

The growth of the urban population around the world has created serious challenges for municipal governments in both developed and less developed countries: air and water pollution, transportation gridlock, deteriorating infrastructure, violence and crime, and income polarization. These problems are particularly acute in developing countries where the population is poor and resources are limited.

Local governments in large metropolitan areas are required to provide a sophisticated transportation and communications network and services that improve the quality of life of people who live there (parks, recreational facilities, cultural institutions). At the same time, large metropolitan areas attract low-income individuals and households seeking employment opportunities and a wider range of specialized social services, thus requiring higher expenditures on social services, social housing, and public health.

Improving the level of basic service delivery is partly a question of resources but it is also a question of governance. Most metropolitan areas are characterized by multiple municipalities each delivering services within their own jurisdiction. The majority of urban services (transportation, water, solid waste, and housing, for example) spill over municipal boundaries. As the physical expansion of large cities extends into the rural hinterland and crosses into other political jurisdictions, there is an increasing need to coordinate service delivery.

Why Governance is Important

How metropolitan areas are governed (including the role of local and regional governments, civil society, business associations, and non-profit organizations) affects their ability to coordinate service delivery across municipal boundaries. It also affects their capacity to deliver and pay for services, the efficiency with which they can deliver services, and their ability to share costs throughout the region in a fair and efficient way. The governance structure of a metropolitan area also has an impact on citizen access to government (the degree to which citizens are involved in decision-making) and government accountability to citizens (how responsive the city
is to their demands). An effective system of governance for the entire metropolitan region is needed to achieve all of these objectives. Yet, few cities in the developing world (or the developed world) have an administrative structure at the metropolitan level.

**Outline of the Paper**

This paper reviews metropolitan structures in many cities around the world. The first part outlines a number of criteria that are used to evaluate metropolitan governance structures including efficiency (with emphasis on economies of scale and externalities), equity, accessibility and accountability, and local responsiveness. The second part reviews different models of government structure: one-tier fragmented government model, one-tier consolidated government model, two-tier government model, voluntary cooperation, special purpose districts, and the role for senior levels of government. The third part sets out eight observations on governance in metropolitan cities around the world. The fourth part offers some concluding comments.

**Findings on Metropolitan Governance**

The findings of this review suggest that neither theory nor practice tells us clearly which model of governance is best for large metropolitan areas. The appropriate model depends on both the national and local context. Models that work well in one jurisdiction may or may not be appropriate in another jurisdiction. There is no “one size fits all” model of metropolitan governance.

Other findings from the review of case studies in many cities in developed and less developed countries highlight the following:

- Not only are different models appropriate in different jurisdictions, the model in any one jurisdiction can and will change over time to reflect population growth, changes in the economy, or changes in the political situation. Examples of changes in governance models over time include Toronto, London, Cape Town, and Abidjan.
• Although voluntary cooperation may not be the best model in terms of efficiency or accountability, it may be more achievable than a full-scale metropolitan government in circumstances where a metropolitan area is too big to be acceptable as a political or administrative unit or where local autonomy is paramount and prevents a consolidation. Examples of voluntary cooperation include, for example, New York City and São Paolo.

• The governance structure should cover the entire economic region but this is rarely the case. Some exceptions are Cape Town, Madrid, and Stuttgart. Over time, economic boundaries evolve and change; political boundaries are much more difficult to modify. For this reason, there will likely be the need for some inter-municipal cooperation across municipal boundaries.

• Consolidation does not necessarily reduce costs. Toronto provides an example of an amalgamation where the harmonization of wages and salaries as well as the harmonization of service levels resulted in cost increases rather than cost reductions.

• It is important to understand the national and local context of the metropolitan area. For example, increased powers and responsibilities were granted to municipalities in recent years in countries such as the Philippines and Brazil. There have been constitutional reforms in countries such as South Africa, Brazil, and India that have affected the role of local governments. The context will have a bearing on what is achievable. Where there is a strong tradition of local autonomy, such as in the Philippines for example, it will be more difficult to implement a metropolitan-wide structure that requires cooperation from local governments.

• The process of reforming the government structure is often as important as the outcome itself. All of the relevant stakeholders need to be included in the early stages of restructuring reform for the outcome to be accepted. Stuttgart, for example, is not widely supported by the municipalities in the region in part because of the way in which it was implemented.
• Politics (and not economics) often dictate the ultimate structure. The criteria of efficiency and equity are not necessarily considered in designing a new governance structure. Both London (with the abolition of the Greater London Council in 1986) and Toronto (with the 1998 amalgamation) provide examples of cities where politics dictated the outcome.

• Metropolitan governments need sufficient resources and fiscal autonomy to raise the funds they need to deliver services, but they rarely have them. For example, the Greater London Authority has limited local autonomy and the City of Toronto funds some social service costs from property tax revenues. Manila is highly dependent on central government grants.

In conclusion, few problems and processes stop at municipal boundaries and the most feasible solutions require larger geographical units and access to a larger pool of resources, both human and financial, than is likely to be at the disposal of small local governments. A strong regional structure that encompasses the entire economic region is essential to ensure that services are delivered in a coordinated fashion across municipal boundaries.
Interest in metropolitan cities around the world is on the rise, in large part, because more and more people are living in cities.¹ The urban population in more developed regions is projected to reach one billion (or 82 percent of the total population) in 2030; the urban population in less developed regions is projected to increase to 3.9 billion people (or 56 percent of the total population) in 2030 (United Nations 2006). Moreover, the number of mega-cities (cities with more than 10 million people) is on the rise. Whereas in 1950, there were only two mega-cities (New York and Tokyo), there were 20 mega-cities in 2005 and the number is projected to increase to 22 by 2015. Developing countries will have 17 of the 22 mega-cities in 2015 (United Nations 2006).

The growth of the urban population has created serious challenges for municipal governments in both developed and less developed countries: air and water pollution, transportation gridlock, deteriorating infrastructure, violence and crime, and income polarization. Moreover, the magnitude and complexity of local government expenditures in large metropolitan areas is greater than in smaller urban or rural areas for many reasons, such as the size and concentration of the population (Freire 2001) and the presence of a more socially and economically heterogeneous population (Nowlan 1994). Local governments in large metropolitan areas are required to provide a sophisticated transportation and communications network and services that improve the quality of life of people who live there (parks, recreational facilities, cultural institutions). At the same time, large metropolitan areas attract low-income individuals and households seeking employment opportunities and a wider range of specialized social services, thus requiring higher expenditures on social services, social housing, and public health.

In addition to providing varied goods and services in greater quantities, metropolitan governments must manage a more complex urban environment; structure land use to promote

¹ The literature on metropolitan governance uses a number of different terms to describe metropolitan areas – metropolitan cities, metropolitan regions, city-regions, urban regions, to name a few. Generally, these terms refer to cities with a large urban core plus adjacent urban and rural areas that are integrated socially and economically with the core. See Stren and Cameron (2005) for a discussion of how the various terms are applied in different parts of the world.
efficiency; and employ financial tools in a way that promotes efficiency (Bird and Slack 2004a). These tasks are often more complex in metropolitan regions than for local governments in general.

Improving the level of basic service delivery is partly a question of resources but it is also a question of governance and allocation (Stren 2007). Most metropolitan areas are characterized by multiple municipalities each delivering services within their own jurisdiction. The majority of urban services (transportation, water, solid waste, and housing, for example) spill over these municipal boundaries. In transportation, for example, there is a need to coordinate services across municipal jurisdictions. For water, there is a need to determine where treatment facilities will be located. For solid waste, there is a need to determine where garbage disposal sites will be located. With respect to police protection, crime does not respect municipal boundaries so service coordination across boundaries is needed. Moreover, there is a need to coordinate policing with other policies and services. For social services, health services, and education, there is a need to determine the appropriate level of expenditures and how costs will be shared among municipalities in the region.

Local governance (which includes, for example, local and regional governments, civil society, business associations, and non-profit organizations) affects the quantity and quality of services, the efficiency with which services are provided, and whether costs are shared throughout the region in a fair and efficient way. The governance structure of a metropolitan area also has an impact on citizen access to government (the degree to which citizens are involved in decision-making) and government accountability to citizens (how responsive the city is to their demands).

This paper provides some observations on the governance of metropolitan areas with some examples from case studies from around the world. It is not intended to provide a comprehensive review of each of the case studies but rather to use them to illustrate some of the observations about urban governance that appear in the literature. The case studies were chosen to reflect

---

2 In one of the few empirical studies of governance and urban performance, Kaufmann et al. (2004) construct a worldwide database of cities containing some key determinants of city performance. They find that good governance (and globalization) at both the country and city level matter for city level performance in terms of access to services and quality of delivery of infrastructure services.
different models of metropolitan government in different regions; they were also selected on the basis of the available literature on municipal restructuring. A key observation from these case studies is that different models have been used at different times in different parts of the world; there is not one model that stands above the rest. The type of model that is appropriate in any city is context-specific. It depends on the responsibilities of municipalities, the revenue sources available to them, the history and culture of the metropolitan area, the intergovernmental context, and other factors.

The outline of the paper is as follows:

- The first part outlines a number of criteria that are used to evaluate alternative metropolitan governance structures.
- The second part describes the types of models of governance structure used around the world. The evaluation of these models considers how well they achieve coordination of service delivery over the entire metropolitan area, the extent to which they allow for the equitable sharing of costs of services throughout the region, and their ability to reduce negative or positive spillovers of service delivery across local boundaries.
- The third part sets out eight observations about governance structures based on case studies from several cities in developed and less developed countries.
- The fourth part provides some concluding comments.

1. Criteria to Evaluate Different Models of Government Structure

Several criteria to design government structure have been set out in the public finance literature (Bird and Slack 2004a). These criteria include efficiency (including the ability to achieve economies of scale and the ability to reduce negative externalities across local boundaries), equity, accessibility and accountability, and local responsiveness. This section sets out the criteria and discusses their application in less developed countries.

3 The UNCHS (1999) lists several indicators that measure urban governance. These indicators provide more specific illustrations of the criteria set out above. The top 12 indicators are: consumer satisfaction (measured through surveys and/or the number of complaints; openness of procedures for contracts and tenders for municipal services; equity in
1.1 Economies of Scale

Economies of scale occur where the per-unit cost of producing a particular service falls as the quantity of the service provided increases. Although this criterion points to the need for large government units, there are problems with its application. First, each urban service will likely achieve the lowest per-unit cost at a different scale of production. For example, the optimal size of government may be different for fire services than for waste management. These differences mean that it can be extremely difficult to draw boundaries for general-purpose local governments based on the criterion of economies of scale.

Second, there is some evidence that larger units of government will result in higher costs for some services because of problems delivering services to remote areas within large jurisdictions or because of “bureaucratic congestion” (Boyne 1992, 336). Governments can become so large that there are diseconomies of scale in the provision of some services.

Third, the jurisdiction that provides the service does not necessarily have to be the one that consumes it. The producing jurisdiction, for example, could sell output to consumers in adjacent jurisdictions. In this way, the producing jurisdiction could benefit from economies of scale in production without having to be part of a larger jurisdiction, that is, without requiring the larger population to be located within its own boundaries. Thus, a larger government jurisdiction is not necessary to achieve economies of scale because the demand and supply of local government services can be separated. Economies of scale can thus be achieved even in a fragmented government system through inter-municipal cooperation or through the creation of special districts to deliver services.

---

Footnote:

4 Economies of scale can be achieved for some services such as water and electricity although most of the economies are achieved in the production of services rather than the distribution to users. For other services such as education, however, economies of scale are difficult to achieve. If the target ratio of students to teachers is 25 to 1, for example, an increase in the number of students will ultimately require more teachers and new classrooms. For a summary of the literature on size economies, see Fox and Gurley (2006).
Fourth, in the context of less developed countries, the impact of a weak infrastructure may negate the advantages of economies of scale. For example, economies of scale may be achieved by having one large school instead of several smaller schools scattered throughout the metropolitan area. If the transportation system is inadequate, however, students may not be able to get to that school. Similarly, citizens may have difficulty getting to a central administration building. Even though there may be economies of scale in centralizing administrative functions, it may still be necessary to decentralize these administrative centers so that people have access to them.

1.2 Externalities

The provision of some services results in externalities (spillovers) whereby the benefits (or costs) of a specific service in one local government jurisdiction spill over on to residents of another jurisdiction. For example, a road in one municipality can provide benefits to residents of neighboring municipalities who also drive on it. In this case of an external benefit, the local government of the municipality in which the road is located has no incentive to provide services to residents of other jurisdictions and is thus unlikely to take account of the external benefits when deciding how much to invest in the road. The result is an under-supply of the service that generates an external benefit. The problem of externalities is one that is often faced by central cities who have to bear the costs of public services to residents of neighboring communities without receiving the related fiscal revenues (OECD 2006, 158).

One way to remove the resulting inefficiency is to design government jurisdictions large enough so that all of the benefits from a particular public service are enjoyed within the boundaries of that jurisdiction. Such boundary readjustments would “internalize” the externalities (ensuring that those who benefit from the service also pay for it).

As with economies of scale, the optimal sized jurisdiction will be different for different services. Furthermore, the optimal jurisdiction from the point of view of internalizing externalities may conflict with the optimal size required to achieve economies of scale. Other ways to address
externalities include intergovernmental transfers and voluntary cooperation among municipalities.

1.3 Equity

Equity refers to the ability to share costs and benefits of services fairly across the metropolitan area. When there are many local government jurisdictions in a metropolitan area, there are likely to be some rich communities and some poor communities. In these circumstances, the rich communities will have a more adequate tax base with which to provide services and may not have very great demands for some services (such as education or social services). The poor communities, on the other hand, may require more services but have only a small tax base on which to levy taxes. The more municipalities within a metropolitan area, the greater will be this problem.

One solution is to consolidate the two (or more) areas into one jurisdiction, in effect taxing the rich municipalities and using some of the proceeds to subsidize the poor municipalities. An alternative approach is to shift the redistributive function to a senior level of government or for the senior level of government to provide transfers to municipalities based on need and fiscal capacity.

1.4 Accessibility and Accountability

This criterion suggests that citizens should have access to local government so that they can influence government policy. Access to policy decisions is particularly important in those countries where democratic traditions are not well established. Access to decision-making can be provided through public meetings, hearings, elections, and direct contacts with officials (Bish 2001). Smaller government units can provide the average citizen with greater access to local government.

5 The transfers would have to be conditional to ensure the funds are spent on the service that generates the externality. They would also have to be matching (that is with some portion of the contribution coming from the local government) to reflect the extent of the externality. For more information on intergovernmental transfers, see Bird and Slack (1993).
decisions because the ability of the public to monitor the behavior of decision makers falls as the size of the government increases (Boyne 1992).

Accountability is closely related to access: the more accessible politicians are to their constituents, the more easily they can be held accountable for their actions. A more fragmented system of local governments should increase public scrutiny and accountability and result in lower service costs simply because the local governments cover a smaller population. The danger of corruption may be greater where there are smaller government units, however, because those who are allocating the benefits are closer to the people receiving them (Lefèvre 2003).

Accountability also requires a link between expenditure and revenue decisions: the body making the decisions about how much to spend should be responsible for raising a large portion of the revenues it requires: “the costs of local decisions should be fully borne by those who make them” (Bird 2001, 117). If there is no accountability in decision-making, there is no incentive to allocate resources efficiently across the different services. Local governments must also be accountable to the central government to the extent that they receive transfers from them.

1.5 Local Responsiveness

The efficient provision of services requires that decision-making be carried out by the level of government that is closest to the individual citizen. This is known as the “subsidiarity principle” and is needed for the efficient allocation of resources, accountability, and responsiveness. As long as there are local differences in tastes and costs, there are clear efficiency gains from delivering services at the local level. According to this principle, expenditure responsibilities should only be assigned to a higher level of government if it can be demonstrated that it can carry out the function more efficiently than the lower level. With few exceptions (such as national defense and services that involve redistribution), almost all public services should be

---

6 The subsidiarity principle was included in the Treaty of the European Union in 1992 in the context of the division of powers and responsibilities between European governmental bodies and their member countries. The principle has also been applied to the role and structure of government at all levels (Barnett 1997).
provided at the local level with local policy-makers making decisions about what services to provide, how much to provide, and who should pay for them.

Public choice theory argues that small-scale, fragmented local governments have special advantages for local democracy because they maintain a quasi-market. The proliferation of small government units in a metropolitan area results in competition among them. Tiebout (1956), for example, suggested that people "vote with their feet," meaning that they move to the jurisdiction with the tax and expenditure package that most closely resembles what they want. Competition benefits citizens through increased efficiency in service delivery or in terms of finding the municipality that has the basket of goods and services that most closely meets their tastes (Boyne 1992). In this framework, a large urban government will be less efficient in meeting the demands of its residents because it will tend to provide a uniform level of public services to people who have different preferences for those services.

The Tiebout model is based on a long list of assumptions. One of those assumptions is that the cost of mobility is zero. Particularly in less developed countries, the ability of people to move around the region for better services and lower tax rates is extremely limited.\(^7\)

1.6 Summary of Criteria for Designing Local Government

The optimal design of government structure depends on which criteria are to be satisfied. Three criteria (economies of scale, externalities, and equity) lend themselves to large government units over an entire metropolitan area; other criteria (local responsiveness and accessibility and accountability) point towards smaller government units. The challenge is to find the right balance between those criteria that are based on economic efficiency and those criteria that are based on responsiveness and accountability. As will be shown in this paper, this balance may be different in different metropolitan areas.

\(^7\) Swianiewicz (2002) stresses this point for Central and Eastern European countries where the ability to migrate in response to variation in local taxes is limited.
2. Models of Government Structure

A variety of metropolitan governance institutions exist around the world. This section reviews four governance models — one-tier fragmented government model, one-tier consolidated government model, two-tier government model, and voluntary cooperation (including inter-municipal agreements and special districts).\(^8\) It also considers the role of senior levels of government in the provision of services.

2.1 One-Tier Fragmented Government Model

In a one-tier fragmented government model, a metropolitan area has a large number of autonomous local government units or special purpose bodies each delivering services within their own boundaries. In some cases, fragmentation is the result of key local functions being managed or delivered by different levels of government (for example, central, state/provincial, and local). Fragmentation leads to poor coordination among the various government units (Montgomery et al. 2003, 365).

Fragmented one-tier governments are common in the United States where metropolitan areas have been characterized by fragmentation, decentralization, and income polarization (Orfield 1997). Klink (2002, 11) describes metropolitan institutions in the U.S. as a large number of “spatially limited networks of general purpose government, special purpose bodies and sectoral institutional structures.”\(^9\) Houston, Texas is an example of a model of “fragmented single tiers” (Savitch and Kantor 2002). Houston is a city surrounded by 790 governments and special districts. These jurisdictions frequently overlap and compete for industry. The Chicago

---

\(^8\) According to the categorization used by Lefèvre (2003), all four models are examples of “governance through institutional building.” Lefèvre refers to the two-tier model as the “metropolitan” model, voluntary cooperation as “metropolitan-wide inter-municipal joint authorities” and special purpose districts as “mono-sector inter-municipal joint authorities.” Conversely, Klink (2002) categorizes models of government structure as “supra-municipal” (two-tier governments) and “inter-municipal” (which includes both voluntary cooperation and special purpose districts).

\(^9\) One notable exception is Portland, Oregon which has a directly elected regional government across three counties and 25 cities. The regional government oversees transportation as well as land use planning, parks and open spaces, civic facilities and the urban growth boundary (Katz 2006).
Metropolitan Area has 464 local governments, including counties, municipalities, and townships (Katz 2006). Because governance in most metropolitan areas in the U.S. is very fragmented, there is rarely a civic or government forum in which to discuss metropolitan issues (Katz 1998).

Most metropolitan cities in the less developed world are also characterized by fragmentation. Mumbai, for example, is a fragmented one-tier structure. The city has a population of 10 million people and is divided into 7 wards each with its own municipal officials. Surrounding the city are the western and eastern suburbs which are also divided into wards. Greater Mumbai (including the suburbs) has a population of more than 18 million. Within the Mumbai urban agglomeration, there is the Mumbai Metropolitan Regional Development Authority, the municipal corporations of Greater Mumbai, Kalyan, and New Mumbai, 16 municipal towns, 7 non-municipal urban centers, and 995 villages. As will be noted in section 2.6, the management of the city is also fragmented because the central, state, and local governments all play a role in delivering services in the city.

The advantage of small, fragmented local governments is that they are more accessible, accountable, and responsive to local citizens. The opportunities to address spillovers, achieve economies of scale, or coordinate service delivery across the metropolitan area are limited, however, compared to the consolidated one-tier or two-tier models described below. Fragmentation creates a policy environment in which metropolitan-wide consensus is difficult to achieve in areas such as economic development, environmental quality, social and spatial disparities, equitable funding of services, and quality of public services throughout the region (OECD 2006, 158). Small, fragmented single-tier governments often engage in voluntary cooperation and participate in special purpose districts that cover the metropolitan area, however.

### 2.2 One-Tier Consolidated Government Model

Under the one-tier consolidated government model of urban governance, a single local government is responsible for providing the full range of local services and has a geographic boundary that covers the entire urban area. Large single-tier governments have generally been
formed by *amalgamation* (the merger of two or more lower-tier municipalities within an existing region) or by *annexation* (appropriation of a portion of a municipality by an adjacent municipality).

Although there are few cases of the one-tier consolidated model, especially in the developing world (Montgomery et al. 2003), Shanghai provides an example of a one-tier consolidated government. Shanghai, with a metropolitan population of more than 13 million, is the largest city in China and one of the largest cities in the world. Within the one-tier, there are municipal and district governments and three tiers of management (municipal, district, and sub-district or street office).\(^\text{10}\)

The one-tier consolidated model is rare in U.S. jurisdictions. The exception is Metro Louisville, Kentucky where a city-county merger was approved by referendum in 2000 – the first consolidation in the U.S. in 30 years. Consolidation reduced duplication, increased the size of the tax base, and made Louisville a more efficient place to do business (Katz 2006).

The main advantages that have been cited for one-tier governments include: better service coordination, clearer accountability, more streamlined decision-making, and greater efficiency (Boyne 1992). Large one-tier governments can take advantage of economies of scale in service provision. Municipal amalgamations internalize externalities: for example, rural residents outside of the original municipal boundary have to pay for urban services that they use. Furthermore, there is funding fairness in the provision of services because there is a wider tax base for sharing the costs of services that benefit taxpayers across the region. The larger taxable capacity of the one-tier government increases its ability to borrow and to recover capital and operating costs from user fees (Bahl and Linn 1992).

---

\(^\text{10}\) Shanghai is one of four Chinese cities that have been granted the status of a province; the others are Beijing, Chongqing, and Tianjin. Provincial or state status is not unique to China. Three cities in Germany have citistate status – Berlin, Bremen, and Hamburg – and Delhi has the political status of a union territory (the National Capital Territory of Delhi) with its own legislative assembly. Citistate status generally gives these large cities some additional powers and more fiscal autonomy than other cities in the country.
How successful consolidated one-tier governments have been in practice at achieving accountability and efficiency (in terms of cost savings) is a matter of debate, however. A large-scale one-tier government may reduce access and accountability because the jurisdiction becomes too large and bureaucratic. To overcome this problem, in some cases community councils or committees have been established to address local issues, or satellite offices have been set up across the municipality where people can pay tax bills, apply for building permits, pay fines, or access city services. Such devices may increase accessibility (if not accountability) but they may also offset potential cost savings that might otherwise result from a larger government unit.

The empirical evidence on fragmented versus consolidated local governments in the U.S. indicates that consolidated structures are associated with higher spending (Boyne 1992). Evidence from municipal amalgamations in Canada also suggests that cost savings have proved to be elusive (Sancton 1996; Slack 2000; Fox and Gurley 2006).¹¹

Amalgamation also reduces competition between municipalities, weakening incentives to be efficient, to be responsive to local needs, and to adapt to changing economic conditions. The reduction in competition may reduce efficiency in the delivery of services and again result in higher costs. On the other hand, if some localities could not previously afford to provide an adequate level of service because they did not have adequate resources, amalgamation may allow them to provide a level of service comparable to richer localities in the region.

Ideally, a large consolidated model would have a geographic boundary that covers the entire urban area or city region. In theory, these boundaries move outward as the urbanized area expands; in practice, however, this is seldom the case. Historically, municipal boundaries have been inelastic, leading to fragmentation of the urban region into city and suburban municipalities (Bourne 2003).

¹¹ The issue of cost savings from municipal amalgamations is discussed in section 3.4 below.
2.3 Two-Tier Government Model

The two-tier government model consists of an upper-tier governing body (usually region, district, or metropolitan area) encompassing a fairly large geographic area and lower-tier or area municipalities (including cities, towns, villages, townships etc.). The upper tier provides region-wide services characterized by economies of scale and externalities (such as transportation, land use planning, and solid waste disposal) whereas the lower tiers are responsible for services of a local nature (for example, local parks). In this way, two-tier models help to resolve the conflict among the various criteria for designing government structure -- economies of scale, externalities, and redistribution on the one hand and access and accountability on the other. Redistribuition throughout a city-region is achieved at the upper-tier level through a combination of tax and spending policies. On the tax side, tax rates are generally levied at uniform rates across the region and the contribution of each lower-tier municipality to the upper-tier municipality depends on the size of its tax base. The larger the tax base in any one municipality, the larger is its contribution to the upper-tier government.

On the spending side, the upper-tier government makes expenditures on region-wide services. These expenditures benefit the entire city-region and are not necessarily distributed among the lower-tier municipalities in the same way as the tax revenues are collected. The result is that a uniform tax (property, income, sales, etc.) at the upper-tier level, combined with region-wide expenditures, serves to redistribute resources from the relatively large tax base municipalities to the relatively small tax base municipalities. There will still be differentiation in service levels and tax rates for services provided by lower-tier municipalities.

With two-tier governments, it is necessary to allocate functions among the tiers. The upper tier should be responsible for services that provide region-wide benefits, generate externalities, entail some redistribution, and display economies of scale. Services that provide local benefits should be the responsibility of the lower tier. Table 1 applies public finance criteria to the various public

---

12 See Barlow (1994) for a discussion of centralization and decentralization arguments in metropolitan areas.
services provided at the local level to determine the appropriate level of government to provide them.

Table 1: Allocation of Expenditure Responsibilities in a Two-Tier Model

<table>
<thead>
<tr>
<th>Function</th>
<th>Upper Tier</th>
<th>Lower Tier</th>
<th>Justification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social services:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Welfare assistance</td>
<td>X</td>
<td></td>
<td>Income redistribution; externalities</td>
</tr>
<tr>
<td>Child care services</td>
<td>X</td>
<td></td>
<td>Income redistribution; externalities</td>
</tr>
<tr>
<td>Social housing</td>
<td>X</td>
<td></td>
<td>Income redistribution; economies of scale; externalities</td>
</tr>
<tr>
<td>Public health</td>
<td>X</td>
<td></td>
<td>Income redistribution; economies of scale; externalities</td>
</tr>
<tr>
<td>Land ambulance</td>
<td>X</td>
<td></td>
<td>Economies of scale; externalities</td>
</tr>
<tr>
<td>Roads and bridges</td>
<td>X</td>
<td>X</td>
<td>Local versus regional roads</td>
</tr>
<tr>
<td>Public transit</td>
<td>X</td>
<td></td>
<td>Externalities; economies of scale</td>
</tr>
<tr>
<td>Street lighting</td>
<td></td>
<td>X</td>
<td>No externalities</td>
</tr>
<tr>
<td>Sidewalks</td>
<td></td>
<td>X</td>
<td>No externalities</td>
</tr>
<tr>
<td>Water system</td>
<td>X</td>
<td></td>
<td>Economies of scale</td>
</tr>
<tr>
<td>Sewer system</td>
<td>X</td>
<td></td>
<td>Economies of scale</td>
</tr>
<tr>
<td>Garbage collection</td>
<td>X</td>
<td></td>
<td>Economies of scale; externalities</td>
</tr>
<tr>
<td>Garbage disposal</td>
<td>X</td>
<td></td>
<td>Economies of scale; externalities</td>
</tr>
<tr>
<td>Police protection</td>
<td>X</td>
<td></td>
<td>Externalities; economies of scale</td>
</tr>
<tr>
<td>Fire suppression</td>
<td></td>
<td>X</td>
<td>Local responsiveness; scale economies for specialized services</td>
</tr>
<tr>
<td>Fire prevention/training</td>
<td>X</td>
<td></td>
<td>Economies of scale</td>
</tr>
<tr>
<td>Local land use planning</td>
<td></td>
<td>X</td>
<td>Local access, responsiveness</td>
</tr>
<tr>
<td>Regional land use planning</td>
<td>X</td>
<td></td>
<td>Externalities</td>
</tr>
<tr>
<td>Economic development</td>
<td>X</td>
<td></td>
<td>Externalities</td>
</tr>
<tr>
<td>Parks and recreation</td>
<td></td>
<td>X</td>
<td>Local responsiveness</td>
</tr>
<tr>
<td>Libraries</td>
<td></td>
<td>X</td>
<td>Local responsiveness</td>
</tr>
</tbody>
</table>

Source: Bird and Slack (2004a)

The Comunidad Autonoma de Madrid (CAM) is an example of a two-tier system with 179 lower-tier municipalities. CAM was created in 1983 and has a population of 5.2 million people. It is administered by a directly elected regional council; the president is elected by the council members. The region took over the previous powers of the Province of Madrid; the lower tiers have fewer powers and responsibilities. The responsibilities of CAM include: transport and infrastructure, education, health, planning, economic development, environment, and culture and research (Lefèvre 2003).
There are a few other examples of two-tier systems at the local level around the world, notably Toronto, Canada (which was a two-tier system from 1954 to 1998); London, England (which recently returned to a two-tier system); Cape Town, South Africa (until 2000); and Abidjan, Cote d’Ivoire (until 2001). These case studies are described below.

Two-tier systems have potentially important advantages in terms of accountability, efficiency, and local responsiveness. Critics of the two-tier model, however, argue that costs are higher because of waste and duplication in the provision of services by two levels of government. Furthermore, two-tier levels of government are less transparent and more confusing to taxpayers who cannot figure out who is responsible for what services. Finally, two municipal councils covering the same geographic area can lead to considerable “wrangling, inefficient decision-making, and delays in implementing policies” (Kitchen 2002).

2.4 Voluntary Cooperation

Voluntary cooperation has been described as "minimal" government restructuring in which there is an "area-wide body based on voluntary cooperation between existing units of local government in the agglomeration with no permanent, independent institutional status" (Sharpe 1995, 12).\(^{13}\) This model is very common in the U.S. and France. Voluntary cooperation is popular, in part, because the area-wide bodies are easy to create politically and can also be disbanded easily. Voluntary cooperation is also common where local autonomy is highly valued: municipalities can retain independence while reaping the benefits of cooperation.

Bologna, Italy, is an example of an inter-municipal model of metropolitan governance on a voluntary basis (Klink 2002). In 1994, forty-eight municipalities and the province of Bologna signed the Accordo per la Città Metropolitana (ACM). The Metropolitan Conference is composed of all the mayors and is presided over by the president of the province. Each

\(^{13}\) Lefèvre (2003) refers to this model in the U.S. and France, as “metropolitan-wide inter-municipal joint authorities,” while Klink (2002) describes it as an “inter-municipal” model.
municipality is free to withdraw at any time and may participate as it chooses in some or all of
the activities of the Metropolitan Conference.

In other countries, voluntary cooperation may take the form of consortia, communities of
communes, urban communities (France), joint inter-municipal authorities (Spain and Belgium),
public bodies, joint agency and core cities (the Netherlands) (Hermann et al. 1999; Jouve 2005;
Klink 2002; Lefèvre 2003; OECD 2006). These forms of cooperation include administrative and
political integration in that there is some form of representation on the boards from the member
local governments. These organizations can levy taxes or collect contributions from the
municipalities or they can levy user fees to pay for services.

In France, for example, nearly all French urban areas have been administered by communautés
urbaines (inter-municipal cooperation institutions for urban areas with over 500,000 people) or
by communautés d’agglomération (inter-municipal cooperation bodies for urban areas with more
than 50,000 people grouped around a central city with at least 15,000 people) since 2003 (OECD
2006; Lefèvre 2003; Klink 2002).

Marseilles provides an interesting example of a move from a model of voluntary cooperation to
two-tier government in 2000 (Klink 2002). The municipalities of Marseilles, Marignane, and
Saint Victoret created a public corporation (the Communauté de Communes Marseilles Provence
Métropole) in 1992 which focused on a few minor projects such as roads and traffic (Savitch and
Kantor 2002). In the following year, thirteen other cities joined this consortium of municipalities
and four more joined in 1998-99. In 2000, the Communauté Urbaine of Marseilles (a
metropolitan organization comprising eighteen cities and one million people) was created. The
regional body (comprised of the mayors and councilors of the constituent municipalities) is
responsible for regional economic development, transportation, land use and housing, crime
prevention, waste disposal and environmental policies. The localities within the Communauté
have adopted tax-sharing agreements whereby the Communauté Urbaine collects a common tax
on business, thereby eliminating tax competition among the local municipalities.
Voluntary cooperation is an alternative way of providing services across a region without resorting to amalgamation. Voluntary cooperation preserves local autonomy, diversity, and the distinct identity of member municipalities (OECD 2006). Municipalities can retain their autonomy with respect to expenditure and tax decisions but, at the same time, achieve economies of scale in service delivery and address externalities associated with service provision (Sharpe 1995). There can be problems of accountability, however, when services are provided by another jurisdiction. Redistribution throughout the metropolitan area is not automatic in a system of voluntary cooperation but could be agreed upon by the municipalities involved. Although these “lighter and more informal forms of governance” can mobilize metropolitan-wide stakeholders around a common vision, implementation requires an action plan and adequate resources that might need a more formal arena for collaboration (OECD 2006).

Notwithstanding the weakness of voluntary cooperation, this form of local governance has steadily grown around the world. One explanation is that voluntarism “is incremental, non-threatening, and capable of growing by trial and error” (Savitch and Kantor 2002). The voluntary model can work well when policy objectives are shared by all policy-makers in the various local governments. Thus, there would be no need for any additional institutional arrangements. It may not work so well, however, when there are divergent objectives. Cooperation usually involves bargaining and some municipalities may not have anything to bargain with. The problems faced by metropolitan areas are significant –global competition, fiscal disparities, urban sprawl – and the solutions may require them to rely on a structure that has a permanent institutional status.

Inter-municipal agreements are formal or informal agreements between municipalities to provide services. They are a type of voluntary cooperation but are less structured in that an official area-wide body is not generally set up to oversee the arrangements. An example of an inter-municipal agreement is the contract services plan in Los Angeles where Los Angeles County provides some services on behalf of municipalities in the Los Angeles metropolitan area on a contract basis. A city-county link occurs in other U.S. jurisdictions as well (Sharpe 1995).

These types of agreements have generally been effective for services such as fire fighting and emergency dispatch, maintenance of boundary roads, purchasing in bulk, and issuing debentures.
Agreements are generally entered into as a way of reducing costs or to set out joint obligations for different municipalities.

Although inter-municipal agreements are successful in achieving coordination and efficiencies for specific services, they are not suitable for achieving region-wide coordination. Furthermore, inter-municipal agreements provide no accountability except through the contract or agreement. If something goes wrong, it is difficult for citizens to know where to complain. Inter-municipal agreements have been described as second-best solutions to reorganization that can lead to "an impenetrable jungle of ad hoc commissions and complex arrangements that even the most conscientious municipal voter will never understand" (Sancton 1993, 33-34).

2.5 Special Purpose Districts

Special purpose districts to deliver services that spill over municipal boundaries provide another way to deliver services in a metropolitan area. Single-purpose special districts provide similar municipal services for several municipalities or manage regional services with externalities. This form of cooperation among municipalities for region-wide services is used in countries where there is a history of strong and autonomous local governments. In the U.S., for example, one third of local governments are special districts or school districts providing education, transportation, water and waste management, economic development, and other services. Joint boards of the special districts are responsible for the management of these services as well as taxing, price setting, and other policy-making. These districts are indirectly controlled by the individual municipal councils.

One of the advantages of special purpose districts is that each service spillover can be addressed on an individual basis. Since it is unlikely that the spillover boundaries are the same for each service, separate districts could be established such as a region-wide transit district or hospital district. Other advantages include: the delivery of services by professionals with decision-making somewhat removed from political influence; services can be provided using more professional expertise than may be available to the municipal government; and dedicated revenues from user fees could be used to finance capital expenditures (Bahl and Linn 1992).
Several problems with special purpose bodies have been identified. First, each body has responsibility for a single service and is not required to make the tradeoffs between, for example, expenditures on transit and expenditures on water and sewers. Second, the proliferation of decision-making bodies has "created a diffuseness of government organizations that is difficult for citizens to understand" (Kitchen 1993). There is no citizen control and confused accountability. Third, there is no direct link between the expenditure decisions made by the special purpose agencies and the local council which collects taxes to fund them. The absence of a link between expenditures and revenues reduces accountability. Fourth, where accountability is lacking, there is no incentive to be efficient. Fifth, when there is a large number of independent special purpose bodies, it is difficult to coordinate interrelated activities and it increases the risk of constituencies emerging to defend sector interests (OECD 2006, 170). Although special districts may work to coordinate individual services across municipal boundaries, they are not suitable for achieving overall regional coordination.

2.6 Role of Senior Levels of Government

Another option to meet the criteria for local government structure is for the national or provincial/state governments to take over the provision of local services. For example, a senior level of government could take over functions such as regional planning and regional economic development. They could also facilitate inter-municipal agreements to improve the coordination of services such as water, waste management, and transit. This coordination function could be done through a national or provincial/state ministry or department. In Australia, a number of “local” services (such as transportation) are provided on a metropolitan-wide basis by the state government.

---

14 Three ways have been suggested to address the problems of coordination. The first is to have overlapping membership so that some of the same people are on a number of district boards. The second is to encourage districts with multi-functions instead of single-purpose districts. The third is to control the operations of the districts so that they remain separate authorities but are still subject to political considerations in the decision-making process. See Bahl and Linn (1992).
The state (and national) government in Mumbai also plays a significant role. The city’s management is divided among the Municipal Corporation of Greater Mumbai, the state of Maharashtra, and the government of India (Patel 2007). The rail network is run by the Railway Board of the Indian Railways. The state government is very much involved in economic development, land policy, housing, and law and order and has established autonomous public sector corporations such as for housing (Patel 2007).

Although provincial/state or national takeover of regional services may effectively address the provision of services that exhibit externalities, the provision of services by a senior level of government also raises concerns about local responsiveness. It may be less appropriate for a senior level of government to provide services because it is further removed from local residents, making it difficult to determine the quality and quantity of output to provide in each municipality. Senior levels of government are likely to be less responsive and less accountable to local residents than a local government.

3. Observations on Governance in Metropolitan Cities around the World

A review of metropolitan cities around the world reveals that there is no one model that stands above the rest. This finding is consistent with other reviews of metropolitan governance (see, for example, Lefèvre 2003; Divay and Wolfe 2002; Bird and Slack 2004a; OECD 2006). The appropriate structure for metropolitan areas depends on several factors that include the roles and responsibilities of local governments, their sources of revenue, service delivery conditions in terms of geography and topography, the homogeneity of preferences in the region, the availability of skilled workers, the political strength of local leaders, the variability in service delivery across the country, the relationship with other levels of government, and other factors (Fox and Gurley 2006).

Although there is not one model that stands above the rest, a review of different models in different cities around the world reveals a number of important findings about the governance of metropolitan areas:
Not only do different governance models work in different cities, models can evolve over time in any one city.

Voluntary cooperation may be necessary where the metropolitan area is too large for a political structure.

Not all consolidated cities cover the entire metropolitan region.

Consolidation does not necessarily reduce costs even though cost savings are sometimes used as the rationale for consolidation.

The national and local context (for example, history and political culture) has a significant influence on how well a given structure will work.

The process of implementing a metropolitan structure is crucial to the success of the outcome.

Politics, and not economics, often dictates the ultimate structure.

For a metropolitan structure to be successful, fiscal resources need to match the expenditure responsibilities but they rarely do.

Each of these findings is discussed below and illustrated with examples from various metropolitan cities in developed and less developed countries.

3.1 Not Only Do Different Governance Models Work in Different Cities, Models Can Evolve over Time in Any One City

Not only have different models of governance been tried in different parts of the world, different models have been tried in the same place at different times. This observation can be illustrated by looking at four cities: Toronto, London, Cape Town, and Abidjan. Although these case studies are all very different – in some of the cities, the restructuring was largely politically motivated; in other cities, it was part of a larger process of decentralization and democratization – the point can still be made that these structures can, and probably should, evolve over time as circumstances change in the metropolitan area. On the other hand, as the Cape Town example shows, too frequent restructuring can result in what Cameron (2005) refers to as “reorganization fatigue.”
Toronto: One-tier to Two-tier to One-Tier

The history of municipal amalgamation in Toronto spans more than 50 years beginning in 1954 with a system of one-tier municipalities, the subsequent creation of a two-tier metropolitan government (a metropolitan tier and 13 lower-tier municipalities) in 1954, and the most recent amalgamation in 1998 which saw the merger of the metropolitan and lower tiers to create a single-tier City of Toronto.

Prior to 1953, the City of Toronto was surrounded by twelve municipalities that had been growing rapidly in the post-war period. The City of Toronto no longer had vacant land for single-family housing development and future growth had to be accommodated in the suburban municipalities. This growth placed huge demands on the suburban municipalities to provide services such as educational facilities, roads, sidewalks, lighting, and sewage disposal. At the same time, most of the suburban municipalities were residential and did not have an adequate tax base to finance the needed services. The City of Toronto, on the other hand, maintained a solid financial base because of its significant commercial and industrial tax base.

In addition to the service demands and inadequate resources in the suburbs, the political boundaries of the City of Toronto no longer reflected the social and economic realities of the metropolitan area (Kulisek and Price 1988, 263). Notwithstanding bilateral agreements with the surrounding municipalities, planning was restricted to the boundaries of the City of Toronto. Further problems arose because each municipality acted independently with respect to transportation, land use, and housing -- issues that really needed to be addressed on a region-wide basis.

In response to these problems, Metropolitan Toronto was formed by provincial legislation on January 1, 1954. Under the Metropolitan Toronto Act, a two-tier government was established with a metropolitan tier encompassing thirteen lower-tier municipalities (the City of Toronto plus the twelve suburban municipalities). The purpose of the two-tier government was to redistribute the wealth of the central city (which no longer had room to grow) to the growing suburbs so that

15 The main source of revenue for Toronto then (and now) was the property tax.
they could provide needed infrastructure; coordinate land use planning and transportation across
the region; and, at the same time, allow lower tiers to be responsive to local needs.

The metropolitan government (Metro) was initially given responsibility for planning, borrowing,
assessment, transportation (transit and some roads) and the administration of justice. Local area
municipalities were responsible for fire protection, garbage collection and disposal, licensing and
inspection, local distribution of hydro-electric power, policing, public health, general welfare
assistance, recreation and community services, and the collection of taxes. Both tiers shared
responsibility for parks, planning, roads and traffic control, sewage disposal, and water supply.
Costs were shared on the basis of property tax base. This cost sharing meant that, in 1954, the
City of Toronto was responsible for 62 percent of the costs of the Metro government.16

Over time, the division of responsibilities between Metro and the lower-tier municipalities
changed somewhat. Metro took over responsibility for police services, social assistance, traffic
control and operations, licensing, conservation, waste disposal, and ambulance services. In 1967,
following the recommendations of a provincially-commissioned report, the number of
municipalities in Metro was reduced from thirteen to six. Property assessment and the
administration of justice became provincial responsibilities in 1970.

Early reviews of Metro government applauded its success in meeting its intended objectives:
“the creation of a federated form of metropolitan government for the city of Toronto and its 12
suburbs in 1953 and the rapidity with which it was able to overcome serious public service
deficiencies made the Toronto model an object of admiration for students of metropolitan affairs
throughout the continent” (Frisken 1993, 153). It provided the necessary infrastructure for the
orderly growth of the suburbs, maintained a vibrant core, and pooled revenues over the whole
metropolitan area. It solved the water and sewage treatment problems, constructed rapid transit
lines, established a network of arterial highways, built housing for seniors, and created a Metro
parks system.

The challenges facing Metro began to change in the 1970s, however. Metro no longer faced growth pressures within its boundaries but, rather, was facing growth outside of Metro in what is now known as the Greater Toronto Area (GTA). Between 1971 and 1975, the province created four regional governments around Metro -- Durham, Halton, Peel and York. It was not until 1988, however, that the province established the Office of the Greater Toronto Area (OGTA) at the provincial level to encourage Metro and the surrounding regions to coordinate their efforts on waste disposal, regional transportation, land use, and infrastructure planning. A forum of GTA mayors (of the lower-tier municipalities) and chairs (of the regional governments) concentrated its efforts on economic development and marketing in the GTA.

The new City of Toronto came into being on January 1, 1998 with the passage of provincial legislation. The upper-tier (metropolitan) government and six local area municipalities were merged into a single-tier city. Following the amalgamation of Toronto, the provincial government also established the Greater Toronto Services Board (GTSB) on January 1, 1999. The GTSB was given no legislative authority except to oversee regional transit. It was not designed to be a level of government nor was it given direct taxing authority. The Board comprised elected representatives from each of the municipalities in the region. The GTSB had limited powers to coordinate decision-making among its member municipalities and to provide strategic growth management. In 2001, after a provincial review, the government disbanded the GTSB. In 2006, the provincial government created the Greater Toronto Transportation Authority (GTTA) to coordinate transportation in a region that includes the Greater Toronto Area plus one additional municipality. It seems that every time a regional authority is disbanded, something else eventually takes its place. The same observation can be made about London below.

**London: Two-tier to One-Tier to Two-Tier**

From 1964 to 1986, London was governed by a two-tier structure: the Greater London Council and 32 boroughs (each with its own mayor and council). In 1986, then Prime Minister Margaret Thatcher abolished the Greater London Council but left the 32 boroughs in place. London’s governance became a direct responsibility of central government ministers (coordinated by a Cabinet sub-committee headed by a Junior Minister for London) and joint agreements. Since
there was no metropolitan authority, ad hoc arrangements were used for regional planning. In
1994, the Government Office for London (GOL) was established to allow the central government
to act as a strategic authority. It brought together the regional offices of line ministries, 32 lower
tiers of local government, the boroughs, and agencies with London responsibilities.

The Greater London Authority Act was proclaimed in 1999 and the new Greater London
Authority with a directly elected Mayor came into being on July 3, 2002. Greater London
comprises 32 boroughs and the Corporation of London. The population of Greater London is 7.4
million.

The new Greater London Authority Act created two new elected bodies – the 25-member
Assembly elected from two different electoral bases (14 on a constituency basis and 11 London-
wide) and the Mayor (who is not a member of the Assembly). Together, the Mayor and the
Assembly constitute the GLA. This system of governance is unique in England.

The GLA’s principal purpose is to promote economic development and wealth creation, social
development, and the environment. It is not permitted to spend directly on any function that is
assigned to the boroughs (such as housing, education, social, or health services). There are four
functions that are separate from the Assembly but accountable to it through the Mayor:

- Transport for London (TFL) is responsible for roads, buses, trains, subways, traffic lights,
  regulation of taxis (metered) and mini-cabs (unmetred and unmarked). The Mayor
  appoints the commissioner. The Mayor chairs the board and appoints 15 non-executive
  members.
- The London Development Agency (LDA) coordinates economic development and
  regeneration. It promotes business and works in partnership with industry, public and
  voluntary sectors. The Mayor appoints the 17-member board and the Chief Executive.
- The Metropolitan Police Authority (MPA) has 23 members of which 12 are Assembly
  members, one is appointed by the Home Secretary, four are magistrates, and six are
  independent Londoners. The police commissioner is appointed by the Queen on advice of
the Home Secretary who shall have regard for any recommendations by the MPA, the Assembly, and the Mayor.

• The London Fire and Emergency Planning Authority (LFEPA) is responsible for fire and emergency services. The Mayor appoints the Chair. There are 17 members of which nine are from the Greater London Authority and eight are nominated by the London boroughs via the Association of London Government.

The boroughs retain primary planning responsibility as the local planning authority. If the Mayor considers an application for a large-scale development to be in contravention of his London-wide strategy, however, he can direct a borough to reject the application. He cannot direct them to approve an application, however.

*Cape Town: No Metro Government to Two-Tier to One-Tier*

Cape Town provides an example of a city that started with no metropolitan government under the system of apartheid, moved to a two-tier metropolitan structure in the mid-1990s, and then to a one-tier metropolitan government with 3.2 million people in 2000. Under apartheid, cities were racially fragmented; only whites had democratic local government structure; black townships had advisory structures (Cameron 2005). Municipal services in white areas were excellent but not so in black areas.

In the 1990s, the interim constitution allowed for three types of local governments – metropolitan, urban, and rural. The Cape Metropolitan Council was established as a metropolitan government. The final constitution in 1996 did not make the same differentiation among types of municipalities. Rather, it provided for three different categories of municipalities. Category A municipalities had exclusive municipal executive and legislative authority in their areas, Category B municipalities shared municipal executive and legislative authority in their areas with a Category C municipality within whose area it fell, and Category C municipalities had municipal executive and legislative authority in an area that included more than one municipality.
Although the African National Congress (ANC) wanted one-tier municipal governments so that the resources and services could be redistributed from former white authorities to former black authorities, they were not able to get there in one stage. A White Paper on Local Government supported single-tier metropolitan government to redress inequalities; promote strategic land use planning and coordinated infrastructure investment; and develop a city-wide framework for economic and social development (rather than compete for investment in uncoordinated fashion). In 1998, despite provincial opposition, two-tier municipalities were replaced by one-tier municipalities which were classified as Category A municipalities.

Preliminary findings on metropolitan government in Cape Town suggest that constant restructuring results in what Cameron (2005) refers to as “reorganization fatigue.” In the last 10 years, Cape Town has seen structural reform (twice), territorial changes (twice), new management structures, and new forms of service delivery. As Cameron (2005) notes, continuous reorganization can disrupt local government and hamper service delivery because so much of the resources and time of the municipal administration are devoted to the restructuring.

*Abidjan: One-tier to Two-tier to One-tier*

Abidjan, with a population of 3.5 million, is the largest city in the former French colony of Côte d’Ivoire. Decentralization began in 1978 with the restoration of “commune” (municipal) status to the major cities in the country including 10 communes in Abidjan. At the same time, the 10 Abidjan communes were amalgamated into a second-tier of government (the City of Abidjan). In 1985, communal status was extended to 98 smaller towns (Stren 2007). In 1989, communal status was extended to the remaining rural areas of the country. The number of local government units increased from 136 to 196. At the time of decentralization, the central government strengthened its advisory and central administrative systems so that it could provide technical support to local governments (Stren 2007).

Until 2001, Abidjan was a two-tier structure with an upper tier and 10 communes each with an elected mayor and councilors and each with differing size and wealth (Montgomery et al. 2003). The upper tier (the City of Abidjan) consisted of the 10 mayors plus four councilors from each
commune. The mayor of the City was indirectly elected by mayors of communes. After his election, he resigned from the communal position at the local level, assigning his position to his assistant mayors.

The major functions of the upper-tier government were waste disposal and management, public lighting, sanitation, traffic regulation, maintenance of roads, parks, and cemeteries, and town planning (Montgomery et al. 2003). The lower tiers (the communes) were responsible for markets, allocation of plots for public purposes, maintenance of primary schools and clinics (but not school or health policy or supervision or payment of professionals), operating social centers, and sharing functions with other levels of government for pollution and hygiene. Private companies were responsible for waste removal, electricity, and water.

Although this system functioned for over 20 years, there were problems (Montgomery et al. 2003). For example, the local government was constrained by the national government in carrying out its assigned functions such as inspection of construction sites and issuing of drivers’ licenses. The City had little direct influence over its finances because the national government collected property taxes and remitted them to the communes which then paid a fixed proportion to the City. There were major differences in the ability of each commune to raise funds. The mayor, because he was indirectly elected, did not have an independent political base from which to promote metropolitan-wide policies.

In 2001, the two-tier government was replaced by the “district” of Abidjan which includes the original 10 communes plus three large sub-prefectures (which include their own communal governments) on the outskirts of the city (Stren 2007). The District is administered by a governor appointed by the president of the republic and assisted by a 51-person district council. The city has expanded in size to include some rural areas. Urban planning is a district level function. As Stren (2007) notes, however, security issues have taken precedence over normal service delivery issues following an attempted military coup in 2002. Service delivery is also constrained by limited local resources.
3.2 Voluntary Cooperation May Be Necessary Where the Metropolitan Area Is Large

Many cities around the world have opted for the model of voluntary cooperation over the more structured model of a consolidated metropolitan government. Voluntary cooperation is increasingly being advocated in metropolitan areas that are evolving in space and time (OECD 2006). Although it may not be the best model in terms of efficiency or accountability, voluntary cooperation can foster communication and possibly limit the tendency to “bureaucratic mission creep” (OECD 2006, 191).

Where a metropolitan area is too big to be acceptable as a political or administrative unit or where local autonomy is paramount and prevents a consolidation, inter-municipal cooperation or special purpose bodies may be more achievable than a full-scale metropolitan government. For example, the New York metropolitan region has never had a general-purpose regional government or a consolidation of local governments because of its deep political and social divisions and strong tradition of home rule. It has, however, had regional authorities (such as the Regional Planning Association) which have acted in the regional interest (Yaro 2000, 75).

Where objectives are shared by policy-makers in various local governments, voluntary cooperation can work to some extent and can be a first step towards a more formal governance collaborative framework (OECD 2006). There are many examples of voluntary cooperation in metropolitan areas in both developed and less developed countries; two examples from less developed countries are described below – one that is considered to be successful (São Paolo) and one that is not (Manila).

*São Paolo ABC Region*

The São Paolo ABC Region is an example of a “bottom up” approach to metropolitan governance for a small part of the São Paolo region. The São Paolo metropolitan region includes the City of São Paolo and 38 surrounding municipalities with a total population of 17.8 million. It is the second largest metropolitan area in Latin America after Mexico City.
Although there is no institution of metropolitan governance for the region, there is an Inter-municipal Consortium of the Greater ABC Region which comprises 7 cities and 2.4 million people. In the early 1990s, the leadership of the region became aware of the severe challenges faced by the economic restructuring that was taking place as well as the inability of existing institutional structures to address these challenges (Klink 2002). The shared nature of the problems helped to forge a new regional identity and led community leaders and politicians to tackle the problem of economic decline through a number of initiatives that included strategic partnerships, integration of fragmented and highly sector strategies, and capacity building (Rodriguez-Pose, Tomaney, and Klink 2001).

In 1990, the Inter-municipal Consortium of the Greater ABC Region was created. Its main focus was to coordinate policies that had spillover effects across municipal boundaries. Although the process of inter-municipal coordination weakened in the mid-90s because the mayors did not consider the regional agenda to be a top priority, the local community itself undertook several initiatives. One of these initiatives was the creation of the Forum on Issues of Citizenship comprising more than 100 NGOs with an emphasis on regional issues. Since 1997, more than 20 agreements on economic, social, and territorial development have been signed.

Seven municipalities in the ABC Region created the ABC Council which is a political body made up of representatives from the State government, the seven local governments, councilors, state legislators from the region, and representatives from civil society. The purpose of the Council is to promote economic development of the region through consensual planning and implementation of innovative public policies (Levy 2002). The City of São Paolo, however, has not participated.

The Regional Development Agency was created in 1998 with a board of directors comprised of private sector members who control 51 percent of the board and the Inter-municipal Consortium with the remaining 49 percent. This Agency is considered to be the legal branch of the Council of the ABC region and can sign technical agreements with external agencies and receive financial resources (Rodriguez-Pose, Tomaney, and Klink 2001).
Although this type of cooperation does not represent a formal structure of governance nor does it include all services or even the entire metropolitan region, some authors have noted that it takes a flexible and pragmatic approach to problem solving. It has operated on the basis of pilot projects that have incrementally built up trust among the main actors (Klink 2002). It needs to find better financing mechanisms, however, and to improve its relationship with the state government and increase the participation of municipal councils (Rezende and Garson 2006; Rodriguez-Pose, Tomaney, and Klink 2000).

**Metropolitan Manila Development Authority (MMDA)**

Metro Manila, with a population of almost 11 million people, incorporates 17 municipalities. The “urban field” of Metro Manila includes another 18 local government units and another four million people (Laquian 2002). The rapid growth in the outlying towns has resulted in physical sprawl and has placed strain on municipal services; coordination and planning at the metro level are weak (Montgomery et al. 2003).

The Metropolitan Manila Development Authority (MMDA), which was created in 1995, is a two-tier metropolitan structure. As required by the constitution, the MMDA was approved by plebiscite. The governing board is the Metro Manila Council (which comprises the 17 mayors of the local government units plus the president of the Manila Vice-Mayors League and the president of the Metro Municipal Councilors League). The Chair heads the council and is assisted by a Deputy Chair, general manager, and assistant general managers for planning, operations, and finance and administration, all of whom are appointed by the president of the Philippines. The MMDA is not a corporate unit of government but rather a special development and administrative unit under the direct supervision of the president of the country (Laquian 2002). It performs planning, monitoring, and coordinating functions but can only perform these functions if they do not diminish the autonomy of local governments on local matters (Laquian 2002).

The MMDA is responsible for services which have a metropolitan-wide impact or that require a level of expenditures that is beyond the capacity of the individual local government units. Its
responsibilities include development planning, land use planning, urban renewal, housing, solid waste disposal and management, development and investment planning, transportation and traffic management, flood control and sewage management, pollution control, and public safety. It derives resources from the central government, a 5 percent contribution from the local government units, and fees and fine revenues.

Criticisms of the MMDA are that it is a national corporation rather than a local institution. Its powers and budget are limited especially when compared to the local authorities whose powers were increased in the 1987 constitution and the 1991 Local Government Code (Montgomery et al. 2003). Moreover, the financial resources from the 17 local governments are often delayed or payments are withheld. There are also issues around having an appointed executive rather than an elected chairman.

Laquian (2002, 79-80) describes the ineffectiveness of the MMDA as follows:

It prepares comprehensive development plans that are neither officially adopted nor followed. Municipalities and cities issue zoning codes and regulations that are not coordinated with the plans. MMDA is supposedly in charge of transport and traffic management but central government controls the financing, construction and maintenance of roads and bridges. It is in charge of garbage disposal but provinces and municipalities will not allow it to set up sanitary landfills or dumps within their territories. It is responsible for urban renewal but housing funds are controlled by the National Housing Authority and other agencies. It has no control over the water system, which has been privatized, nor over the design and construction of rapid transit systems and toll roads, which have also been privatized.

3.3 Not All Consolidated Cities Cover the Entire Metropolitan Region

Even in cases where there has been a consolidation to one large city, it often does not cover the entire metropolitan region (OECD 2006). Even where it does cover the economic region at the time of the consolidation, the economic boundaries will continue to expand over time. Economically dynamic regions, by their nature, eventually outgrow their local political boundaries. Although the boundaries of the economic region grow incrementally over time and space, government boundaries are difficult to change and only change occasionally “in big leaps,
and often after complex procedures” (OECD 2006, 157). As a result, the consolidated city still needs to coordinate services such as transportation and economic development with neighboring municipalities.

In reality, few merged areas cover the entire metropolitan region (OECD 2006). Some brief examples illustrate this point:

- The Greater London Authority (GLA) does not cover the economic region of the South East. The GLA includes 7.5 million people. The economic city, which is the Greater South East, has a population of 20 million.
- The ABC Region of São Paolo does not coincide with the administrative region of São Paolo or the functional region.
- The amalgamated City of Toronto is too small and too big at the same time. The metropolitan government is too small to cover the economic region or address region-wide spillovers related to transportation and planning and it is too big to be locally responsive and accessible. The population of the City of Toronto is 2.5 million; the population of the economic region (the Greater Toronto Area) is over 5 million.

Some cities, however, do cover the economic region.

- The boundaries of Cape Town were drawn by the Municipal Demarcation Board so that 95 percent of the people who live and work in Cape Town also reside there.
- The Comunidad Autonoma de Madrid (CAM), which was created in 1983, covers more or less the functional area of the metropolis. CAM has 5.2 million people and includes 179 municipalities. The Region has taken over previous powers of Province of Madrid.
- The Stuttgart city-region comprises 179 local authorities (including the city of Stuttgart). The Verband Region Stuttgart was established by Land legislation in 1993 as an association of local authorities in the Stuttgart region. The city-region, with a population of 2.6 million, covers more or less the functional territory of the metropolitan area.
3.4 Consolidation Does Not Necessarily Reduce Costs

Although cost savings are often the overriding objective of consolidation, these savings often fail to materialize. There may be some cost savings if economies of scale can be achieved but economies of scale are very service-specific (Fox and Gurley 2006). For example, it is more likely that economies of scale will be achieved for services such as water and sewers than for education. The potential for savings depends on several factors such as the willingness to eliminate redundant workers. Even where there are cost savings from consolidating staff, there is a tendency for wages and salaries to equalize up to the level of the highest expenditure municipality. Similarly, the harmonization of service levels to the highest expenditure municipality results in higher costs. Finally, there will be transition costs in moving to a new government structure.

A review of empirical evidence in the U.S. on fragmented versus consolidated local governments concludes that lower spending is a feature of fragmented local government systems; consolidated structures are associated with higher spending (Boyne 1992). Indeed, as noted earlier, very large consolidated units of government will be subject to “diseconomies of scale” because of “bureaucratic congestion” and problems delivering services to remote areas. Moreover, amalgamation reduces competition between municipalities because there is less incentive to be concerned with efficiency and less incentive to be responsive to local needs. Competition increases efficiency in the delivery of services and results in lower costs.

Toronto

The amalgamation in Toronto provides an example of an amalgamation that did not result in cost savings (Slack 2000). The argument on cost savings asserts that amalgamation can reduce the number of politicians and administrators. Although this argument is valid, it also true that the amalgamation of municipalities with different service levels and different wage scales results in expenditure increases. As Tindal (1996, 50) notes: “past experience tells us that there are strong upward pressures on costs after an amalgamation.”
In the case of Toronto, cost savings were likely to be fairly small since the largest expenditures (welfare assistance, transit, and policing) were Metro responsibilities prior to the amalgamation in 1998. These three services accounted for 70 percent of the total upper-tier and lower-tier expenditures combined. This means that any cost savings that could be achieved only applied to 30 percent of the total budget of the new city.

Most of the savings that did occur resulted from staff reductions. Reports following the amalgamation suggest that between 1998 and 2002, 2,700 positions were eliminated through amalgamation. Over the same time period, however, an additional 3,600 positions were added to improve service levels, thus resulting in a net increase in employment over the period (Schwartz 2004, 101-102). Unfortunately, it is not possible to determine if these new positions were created in response to the amalgamation or to other provincial initiatives that were introduced in the same year.

When municipalities are amalgamated, some duplication is obviously eliminated. In particular, the number of politicians and bureaucrats is reduced. On the other hand, when municipalities with different service levels and different wage scales merge, expenditures may increase. For example, when the fire departments of several municipalities are amalgamated, it is possible to reduce costs by eliminating a number of fire chiefs. However, the newly amalgamated city faces thousands of fire fighters who are doing the same job and working for the same employer – the new city. Understandably, they will want to be paid comparable salaries and benefits, and none will be willing to take a pay cut. Salaries and benefits thus tend to equalize up to the highest expenditure municipality. This upward harmonization of wages and salaries generally outweighs

---

17 Prior to the amalgamation, the provincial government commissioned a study on the potential savings from the Toronto amalgamation – KPMG Consulting (1996). This study, however, estimated cost savings in the area of policing, a service that had already been amalgamated in 1967. The study also did not estimate the potential impact of the harmonization of wages and salaries and services that would occur following the amalgamation.

18 Local services realignment across the province, for example, resulted in the transfer of responsibility for many services such as water, sewers, roads, transit, social housing, public health, ambulances, and increased responsibility for social services to municipalities. In return, the provincial government uploaded primary and secondary education and a few other services to the provincial level.
any cost savings. Amalgamations thus often increase costs by harmonizing service levels at the highest service level enjoyed before the amalgamation.

In addition to the increased costs associated with service and wage harmonization, municipal amalgamations generally result in transitional costs and they are often higher than anticipated (Vojnovic 1998). If the transition is towards a more efficient, effective, and accountable local government, then the costs are probably worth it. If not, transitional costs become an additional argument against amalgamation. One-time transition costs in Toronto, for example, included acquisition of new technology for financial, human resources, and payroll systems; renovation of existing facilities such as the Toronto City Hall; and hiring of technical and professional expertise with respect to areas such as telecommunications (City of Toronto 1999). In addition to the one-time costs, there were also costs associated with staff exits.

3.5 The National and Local Context (for Example, History and Political Culture) Has a Significant Influence on How Well a Given Structure Will Work

The governance of metropolitan areas depends very much on the national and local context in which those areas are operating. In comparing four cities (Johannesburg, Mumbai, São Paolo, and Shanghai), for example, Segbers (2007, 1) notes that they each have different local legacies, resources, and national contexts and these have an impact on the type of governance structure that has emerged in those metropolitan areas.

Three case studies are reviewed below to illustrate how the context matters -- Shanghai, Manila, and São Paolo. In the case of Shanghai, the lack of democracy and the central control over municipalities have an important role to play in the governance of the city. Although making Shanghai a province may have merit, this move would be much more difficult to implement in democratic countries. In the case of Manila, people identified metropolitan government with the Marcos regime and thus opposition to the regime had an impact on their perceptions about

---
19 Of course service delivery costs could be reduced without changing government boundaries, for example, through such methods as the provision of services by one municipality to other municipalities and the involvement of the private sector in service delivery.
metropolitan government. In the case of São Paolo, the presence of a military regime and a new constitution had an impact on the desire for local autonomy in that metropolitan area.

Shanghai

Shanghai is a unified governance structure with authority over local governments in the district. Because the municipality is under direct control of the central government, the local governing authority has considerable power. The Mayor is appointed by the central government and exercises authority delegated to him but can also make some autonomous decisions. He has the authority to initiate projects of $50 million or less without approval from Beijing, for example (Laquian 2005, 113). Shanghai also has the status of a province.

In 1988, as part of the fiscal decentralization, cities in China were given more autonomy in revenue collection and expenditures. Local governments have also gained more discretion over investment and the promotion of local growth (Wu 2007). Other changes include the introduction of the market mechanism to the development and use of urban infrastructure (Wu 2007) and the delivery of water, gas, and electricity by utility companies. User charges are levied for buses, gas supplies, waste water discharge, and municipal sanitation services. Tolls have been used on urban roads, bridges, and tunnels to pay for the investment in the infrastructure.

Institutional reforms have meant that there are two tiers of government (municipal and district) and four tiers of management (municipality, district/county, sub-district/ street office, and residents’ committees). Sub-districts are agencies of the district. Since 1990, the urban districts of Shanghai have gained new administrative powers such as planning, public works maintenance, approval of local foreign trade deals, and commercial administration. Street offices are not a level of government because there is no People’s Congress at that level. Rather they act as a representative or agency of the district government (Wu 2002). They are responsible for policing and basic health care and receive between one and two percent of the of district government tax. In recent years, the role of the street offices has been strengthened and professionalized. Community service centers have been set up to provide property management, community cleaning and sanitation, leisure and recreation, social welfare services; community
centers exist at district and sub-district levels. Residents’ committees are elected by residents and undertake many tasks assigned by the government such as maintenance of public order and basic welfare provision.

The Street Office is considered to be the base level of government in the city with comprehensive regulatory powers and more than 150 daily management functions, including, for example, local social security, employment arrangements, administration of migrant workers, and maintenance of market orders (Wu 2002, 1084). As Wu (2002) notes, the street office is evolving towards a government body. Experiments are being tried to increase base level democracy by establishing residents’ self-governing bodies. The need for this kind of accountability to the population is noted by Bird (2004, 11) who argues that “it is essential to build in more open channels of citizen feedback to create the resilient ‘error-correction mechanisms’ that constitute an essential component of any sustainable political-economic system.”

Manila

Manila provides an example of a city in a country with a long history of local autonomy that has made cooperation at the metropolitan level difficult. Moreover, the affiliation of metropolitan structures in the mind of the public with the Marcos regime has further limited the ability to introduce such a structure in Manila.

The history of metropolitan authorities in Manila goes back to the 1960s when migrants from across the country moved to the city. The rapid growth put strains on the municipal infrastructure and put pressure on the resources of the local government units. The need for a regional coordinating body became evident at that time. The mayors of Manila and the neighboring municipalities organized themselves into a league to address the most pressing problems in the region. Since membership in the league was voluntary, however, it was unable to coordinate long-term development.
The Metro Manila Commission (MMC) was created in 1975 following a successful referendum on a proposal to create a single metropolitan area by integrating four cities and 13 municipalities. Under the MMC, all legislative and executive authority was vested in a small governing body appointed by and under the control of the President of the Philippines. The role of the MMC was to perform general administrative, executive, and policy making functions and to provide services common to the metropolitan area. Funding came from the 17 local government units in the metropolitan area who were required to contribute 20 percent of their annual income to the MMC.

Popular support for the MMC declined in the Marcos years and, in 1990, the new president created the Metro Manila Authority (MMA) to operate as a metro-wide authority in the National Capital Region. The MMA had jurisdiction over basic urban services including land use planning, traffic management, public safety, urban development and renewal, and waste management. It was governed by a metropolitan council comprising the mayors of the 17 local government units and headed by a Chair who was indirectly elected by the members every six months. It continued to collect revenues from the local government units but the amount was reduced to 15 percent of their annual income. The Metropolitan Manila Development Authority was created in 1995.

The Metropolitan Manila Area has been described as “city of villages” with autonomous local units resisting higher level controls of their activities (Laquian 2002, 74). On the basis of the deliberations of council, Laquian (2005, 113) makes the point that it is obvious that most local executives are only concerned about their own political units and not the metropolitan area as a whole. To prove his point, Laquian (2005) describes an incident in 2003 when the chair of the MMDA attempted to improve traffic conditions by experimenting with number coding so that vehicles with plates ending in odd numbers can only be used on odd-numbered days. The mayors of two cities refused to join the experiment which resulted in major traffic jams.

One of the things that stand in the way of the effectiveness of the MMDA is a law which stipulates that its activities should be carried out “without prejudice to the autonomy of the affected local governments.” The MMDA has remained a weak institution because of a need to
preserve local autonomy and, as a result, urban services remain uncoordinated. People have strong loyalties to very local government units (Stren 2002, 7). Politically and administratively, “the citizens of Metro Manila find it easier to identify with their local interests” (Laquian 2002, 84). There has never been a case where two municipalities have merged to meet commonly shared challenges (Laquian 2002).

São Paolo

Metropolitan governance in São Paolo needs to be understood in the context of the military regime and subsequent introduction of a new constitution (Souza 2005). In 1967, during the military regime, the metropolitan issue was incorporated in the constitution and formalized through a federal law in 1973. The framework for metropolitan governance was highly centralized and authoritarian; there was not much scope for the participation of local actors. In the 1980s, economic problems in the country meant that the central government was sending fewer funds to metropolitan regions. There was also resistance to the top down approach being used.

In 1988, the new constitution delegated responsibility for designing metropolitan structures to state legislatures. The constitution recognized municipalities as members of the federation with similar status to states and increased their political, administrative, financial, and legislative autonomy (Levy 2002; Rezende and Garson 2006). Municipalities were assigned control over intra-city transport, pre-school and elementary school education, land use, preventive health care, and historical and cultural preservation (Montgomery et al. 2003). Moreover, municipalities were given the right to establish councils of stakeholders (municipal boards or community councils) and these bodies were established in most of the largest cities (Montgomery et al. 2003). These bodies include non-elected representatives of community groups and address matters such as urban development, education, the environment, health, and sanitation.

Although the states could establish metropolitan governments in Brazil after 1988, the states could not interfere in municipal autonomy. The result is that any institution set up to oversee metropolitan activities is only administrative in nature and has no legislative power. Cooperation
is further exacerbated by a tax sharing formula that results in significant inequality in per capita budgets in metropolitan regions (Rezende and Garson 2006). This inequality reduces the incentive to cooperate because rich municipalities do not want to contribute to investments outside their municipal boundaries and poor municipalities find it easier for their residents to access services in neighboring jurisdictions. Inter-municipal coordination thus depends on the willingness of the individual mayors to cooperate.

According to Klink (2002), the changes in the constitution did not result in major changes in metropolitan governance in Brazil. He argues that the standard criteria for designing government structure (economies of scale, spillovers, etc.) had been used by the military regime in a top down manner and thus there was resistance to applying these criteria in the context of the constitutional assembly. Rather, issues of re-democratization and autonomous local government dominated the policy agenda. The result is a federal system with fragmented local governments. Although there are few examples of inter-municipal cooperation, the ABC region described earlier has played an important role in economic development for at least a part of the São Paolo metropolitan area.

3.6 The Process of Implementing a Metropolitan Structure Is Crucial to the Success of the Outcome

The process followed in getting to a metropolitan structure is critical to its success. As Klink (2002, 27) notes: “the product (reduction of territorial spill-overs, collective goods, economies of scale, delivery of specific metropolitan services etc.) and the process (voice, transparency and accountability) of the metropolitan agenda cannot be separated as if these were two different issues.”

The success of metropolitan governance reform depends, to a large extent, on the public support for the new structure and that support depends, at least in part, on the process followed to establish it (OECD 2006, 196). In the case of the more institutionalized models, success will depend on how well the public has been prepared for the change.
Generally, metropolitan governments are established through a top-down process and local actors have to deal with a structure that they did not ask for and do not want (Lefèvre 1998). In many cases, local actors are not consulted. In other cases, they are consulted but their views are not followed (such as the forced amalgamation in Toronto). In cities where there has been consultation, negative votes have prevented amalgamations from occurring such as in Amsterdam, Rotterdam, and Berlin (Lefèvre 1998).\(^{20}\)

As Klink (2002) notes, international experience tells us that the political economy of changing institutional and financial structures requires the bringing together of stakeholders at the initial stages. Incremental bottom up changes will be more successful than top down proposals. Klink (2002) reviews the European experiments with metropolitan governments in the 1970s and concludes that the failure to implement metropolitan governments on the basis of economic criteria (such as economies of scale, spillovers, and redistribution) resulted from these models being imposed by senior levels of government without the participation of the main stakeholders. The result was that local municipalities, political parties, and pressure groups opposed the metropolitan model from the beginning. The only way these models could survive, in places where they did survive, was under pressure from the central or state government (Lefèvre 1998).

Who are the players? The state needs to be involved because it is the only player that can give legitimacy to the process and the outcomes (Lefèvre 1998). Rarely have metropolitan governance reforms derived from purely local initiatives (OECD 2006). Rather, the state government has generally initiated governance restructuring either by imposing or encouraging it. Stakeholders are needed because the support of economic and social interests is important. The population needs to be involved because, in a democratic system, public support is needed for the metropolitan arrangements to last (Lefèvre 2003).

\(^{20}\) Another interesting example in which the public was consulted was in Montreal. In 2000, the province of Quebec announced that it would merge all of the constituent parts of the two-tier urban communities in Montreal (and some other cities in the province). The opposition party in the provincial legislature promised, if elected, to provide a mechanism for de-mergers. When it was elected in 2003, it introduced a somewhat complicated mechanism for de-merger (see Sancton 2005, 324 for more details). De-merger referenda were held in 2004 and 15 de-mergers were approved in Montreal.
The process also depends on the incentives to amalgamate. Threats or fiscal incentives may encourage cooperation. In countries where municipalities do not have original powers in the constitution, for example, the threat that the state government will consolidate municipalities with or without the support of the municipalities or the public, provides an incentive for municipalities to undertake their own consolidations. Fiscal incentives to amalgamate may also encourage municipalities to restructure, although they do not always work. In Latvia, for example, the central government encouraged municipalities to amalgamate by offering grants equal to 5 percent of their budget (with grants falling to 4 percent and 3 percent in subsequent years). Municipalities did not generally see the benefits of amalgamation, even with the additional grant funding, and opposed amalgamation (Fox and Gurley 2006). From 1998 when the program began until 2003, only 21 municipal governments had been eliminated by amalgamation; the government had hoped to eliminate 440 local governments.

Stuttgart

The Verband Stuttgart Region was established after some controversy (Lefèvre 2003). Stuttgart provides an example of a city-region that is fairly weak, in part because of the process used to set it up. The Stuttgart city-region (established through a provincial law in 1993) comprises 179 local authorities (including the City of Stuttgart) in five counties. Most of the lower tiers were against the creation of the regional tier; the result is a weak metropolitan authority with few responsibilities (Lefèvre 2003). Its main responsibility is the public transport authority of the metropolitan area (representing 88 percent of its total budget). It is also engaged in tourism and regional planning. As Lefèvre (2003) notes, although legislation allows the region to undertake other functions, opposition by local governments has prevented it from doing so. Moreover, although the region raises funds from constituent municipalities, it has not been able to do that either because of political reasons.

The Verband Stuttgart Region was set up by the Land (state) government to address different political and economic issues (Jouve 2005). Since the 1960s, the central city and the outlying areas disagreed over how the costs of services provided by the central city (transportation, hospitals, and other services) should be shared. The outlying municipalities refused to pay for
these metropolitan services (Jouve 2005) leaving the city to bear of all of these costs. Following an economic crisis in the early 1990s, the Stuttgart Chamber of Commerce and Industry linked the issue of reforming local government with the issue of economic development. At the same time that the Chamber was placing demands on the city to reform local government, a coalition government came into being in the Land government.

The creation of the Verband Stuttgart Region was only possible because of a strong state government and a unique political situation that resulted in a coalition between the left and right in the Land Assembly. The creation of a directly-elected metropolitan institution was part of the coalition agreement between the two political parties (Jouve 2005). The resulting structure, however, was fairly weak, in large part because it was given no taxing authority and relied on transfers from the Land government and the constituent municipalities. It was also weak because the constituent municipalities did not support it.

London

London, on the other hand, enjoyed strong public support for the creation of the Greater London Authority. It won in a 3-1 vote in a referendum. Although there was not much interest in its predecessor (the Greater London Council), its abolition in 1986 served to increase support for a metropolitan-wide government. Civil society supported the creation of the GLA and of direct election of the mayor (Jouve 2005). The business community and community groups supported the new structure because it allowed them to deal with one local government with strong political legitimacy instead of a series of smaller local governments (Jouve 2005).

3.7 Politics, and Not Economics, Often Dictate the Ultimate Structure

Several criteria for designing local government structure were set out earlier in this paper. Many of these criteria are based on the economics principles of efficiency and equity. Yet, rarely are these economic principles used to determine the optimal government structure. More often, the structures that are imposed are dictated by politics, and not economics. In London, for example, the abolition of the Greater London Council in 1986 by Prime Minister Thatcher was largely on
political grounds – she was a Conservative prime minister and the Mayor of London was a member of the Labour Party. Ironically, the termination of the GLC did more than any other action of the council to unite Londoners behind it (Pimlott and Rao 2002). In Manila, popular support for the Metro Manila Commission (MMC), and indeed metropolitan governments of any kind in the Philippines, declined significantly in the Marcos years when it was perceived as a political puppet of the administration (Laquian 2005).

Toronto

The amalgamation of Toronto provides an example of an amalgamation that was implemented even though studies of governance structure in the region did not recommend it. The Greater Toronto Area (GTA) Task Force was appointed in 1995 in response to growing concerns about the future economic performance of the Toronto region. The major conclusions on governance in the Task Force report focussed on the need to treat the entire GTA as a single economic unit with a unified economic strategy and to create a new GTA governmental body to deal with GTA-wide environmental and planning issues and share major infrastructure and social costs (GTA Task Force 1996). The report also recommended that local government within the GTA needed to be simplified by eliminating Toronto’s upper tier (Metro) and the four surrounding regional governments and by reducing the number of lower-tier municipalities.

The Who Does What (WDW) Panel, appointed by the provincial government in 1996, also called on the Province to set up a GTA governance structure. It recommended the creation of a Greater Toronto Services Board (GTSB) eliminating the five metropolitan and regional governments and consolidating some of the lower-tier municipalities into strong cities.

Notwithstanding these provincially-commissioned reports on the need to coordinate service delivery between Toronto and the surrounding regions, the provincial government chose to amalgamate the municipalities within Toronto instead. Moreover, referenda held in each of the lower-tier municipalities in Metro Toronto did not support the amalgamation. Of the approximately 36 percent of eligible voters who voted, opposition to the proposed amalgamation ranged from 70 to 81 percent of voters, depending on the municipality. Since, according to the
Canadian constitution, municipalities are creatures of the province, these referenda were not binding. The province can, and did, simply introduce legislation that consolidated the municipalities.

Opposition to the proposed amalgamation came from many different quarters: local municipalities (both inside and outside of Metro Toronto), the opposition parties, citizen organizations, and from within the governing party itself (Stevenson and Gilbert 1999). The major citizen opposition centred on the loss of local identity and reduced access to local government. In the broader context of the GTA, it was felt that amalgamation would result in increased polarization within the region.

3.8 Fiscal Resources Need to Match Expenditure Responsibilities but Rarely Do

The amount of funding available to the metropolitan area is an important determinant of the quantity and quality of services that it will be able to provide. As noted above, there are examples in a number of countries of increased powers and responsibilities at the local level over the last several years. Revenues at the municipal level, however, have not kept pace with the increased expenditure requirements (Montgomery et al. 2003; OECD 2006). Not only do metropolitan areas depend heavily on intergovernmental transfers, their own revenue sources are inadequate. In most countries, municipal own-source revenues are generally based on property taxes and user fees and not the more lucrative taxes such as income, sales, and fuel taxes. In many countries, intergovernmental transfers are not reliable. In poor countries, property taxes are difficult to collect (Bird and Slack 2004b) and most municipalities in developing countries are not permitted to borrow (Montgomery et al. 2003).

Metropolitan areas need to have sufficient resources to meet their expenditure requirements, yet they rarely have them. Large metropolitan areas should have greater fiscal autonomy (greater responsibility and ability to manage local services and to levy taxes and collect revenues) than other urban or rural areas (Bird and Slack 2004a). The ability to self-finance is a critical factor in determining which metropolitan institutions succeed or end up bickering between contending
financial supporters. Fiscal autonomy for a metropolitan area means that it can more easily raise the revenues needed to provide services.

In general, the types of expenditures made by metropolitan governments include: police and fire; water, sewers, garbage; transportation; recreation and culture; social services, health, and housing; and planning and development. In some metropolitan areas, primary and secondary education is also a local responsibility.

The main sources of revenue for local governments are: user fees (for transit, water, recreation, and some social services); taxes (including taxes on property, income, and sales); and intergovernmental transfers (conditional and unconditional). Local governments in most countries are not permitted to borrow to meet operating expenditure requirements. They are, however, often permitted to borrow to make capital expenditures. Nonetheless, local access to capital markets is often heavily restricted in developing countries (Rodden, Litvack and Eskeland 2003). Even in countries with well-developed capital markets, smaller municipalities may only be able to borrow through a financing authority or state/provincial body, in part to reduce borrowing costs by pooling the borrowing requirements of different municipalities.

It is very difficult to find comparable data on municipal revenues and expenditures across countries (and sometimes even within countries). The remainder of this section provides examples from a few cities in developed countries for which municipal public finance data are available.

London

The Greater London Authority has very little local fiscal autonomy. As Table 2 shows, the main responsibilities of the GLA are transport and police; other responsibilities include economic development, planning, and fire and emergency planning. Its revenues sources largely come from central government grants which account for over 80 percent of the revenues of the GLA. Central grants include the distribution of centrally-collected non-domestic rates (non-residential
property taxes) on a per capita basis to local government. Other revenues include local property taxation (the council tax) and user fees (other income).

The same can be said about lack of local fiscal autonomy for the lower tiers (the boroughs) in London. As Table 3 shows, the main expenditures for the boroughs are education, social services, and housing. The main sources of revenue are grants from the central government. The council tax (residential property tax) accounts for only 15 percent of borough revenues.

### Table 2: Municipal Expenditures and Revenues, Greater London Authority, 2006-07

<table>
<thead>
<tr>
<th>Expenditures</th>
<th>%</th>
<th>Revenues</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metropolitan Police Authority</td>
<td>33.9</td>
<td>Specific grants</td>
<td>33.6</td>
</tr>
<tr>
<td>London Fire &amp; Emergency Planning Authority</td>
<td>4.5</td>
<td>Other income</td>
<td>34.9</td>
</tr>
<tr>
<td>Transport for London</td>
<td>55.9</td>
<td>Reserves</td>
<td>1.0</td>
</tr>
<tr>
<td>Greater London Authority</td>
<td>1.4</td>
<td>Police grant</td>
<td>10.4</td>
</tr>
<tr>
<td>London Development Agency</td>
<td>4.4</td>
<td>General GLA grant</td>
<td>0.4</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td>Revenue support grant/non-domestic rates</td>
<td>11.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Council tax</td>
<td>8.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Source: The Greater London Authority’s Consolidated Budget and Component Budgets for 2006-07

### Table 3: Municipal Expenditures and Revenues, London Boroughs, 2005-06

<table>
<thead>
<tr>
<th>Expenditures</th>
<th>%</th>
<th>Revenues</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education</td>
<td>37.8</td>
<td>Interest receipts</td>
<td>1.3</td>
</tr>
<tr>
<td>Highways and transport</td>
<td>1.8</td>
<td>Specific grants outside AEF (Aggregate External Finance)</td>
<td>27.5</td>
</tr>
<tr>
<td>Social services</td>
<td>20.7</td>
<td>Specific grants inside AEF (Aggregate External Finance)</td>
<td>12.5</td>
</tr>
<tr>
<td>Social housing</td>
<td>25.8</td>
<td>Use of reserves</td>
<td>-1.2</td>
</tr>
<tr>
<td>Culture</td>
<td>2.9</td>
<td>Revenue support grant</td>
<td>29.5</td>
</tr>
<tr>
<td>Environment</td>
<td>5.0</td>
<td>Redistributed business rates</td>
<td>14.3</td>
</tr>
<tr>
<td>Planning and development</td>
<td>2.0</td>
<td>Police grant</td>
<td>0.2</td>
</tr>
<tr>
<td>Police</td>
<td>0.5</td>
<td>Other items</td>
<td>0.9</td>
</tr>
<tr>
<td>Courts</td>
<td>0.1</td>
<td>Council tax</td>
<td>15.1</td>
</tr>
<tr>
<td>Central services</td>
<td>2.6</td>
<td>Total revenues</td>
<td>100.0</td>
</tr>
<tr>
<td>Other</td>
<td>0.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total current expenditures</td>
<td>100.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Toronto

Toronto spent over US$7.7 billion in 2005. Although the City is required to balance its operating budget (as are all cities in Canada), it nevertheless faces a mismatch between its expenditure responsibilities and revenue-raising tools. The mismatch occurs because the property tax funds a significant amount of redistributive services such as health, social services, and social housing. These services are not appropriately funded by the property tax but rather should be funded by more progressive taxes such as the income tax.  

Table 4 shows the distribution of operating expenditures and revenues in Toronto in 2005. The main expenditures that the city makes are for transportation (roads and transit), social services and social housing, and water and sewers. The main source of revenue is the property tax which accounts for 41 percent of total revenues. Other major revenue sources are user fees and intergovernmental transfers.

<table>
<thead>
<tr>
<th>Expenditures</th>
<th>%</th>
<th>Revenues</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>General government</td>
<td>5.3</td>
<td>Property and related taxes</td>
<td>41.0</td>
</tr>
<tr>
<td>Fire</td>
<td>4.3</td>
<td>User fees</td>
<td>20.6</td>
</tr>
<tr>
<td>Police</td>
<td>9.9</td>
<td>Federal and provincial transfers</td>
<td>20.1</td>
</tr>
<tr>
<td>Roads and transit</td>
<td>19.2</td>
<td>Licenses and permits</td>
<td>5.9</td>
</tr>
<tr>
<td>Water, sewers and solid waste</td>
<td>16.3</td>
<td>Other revenues</td>
<td>12.4</td>
</tr>
<tr>
<td>Health</td>
<td>4.0</td>
<td>Total</td>
<td>100.0</td>
</tr>
<tr>
<td>Social services and social housing</td>
<td>29.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recreation and culture</td>
<td>8.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Planning and development</td>
<td>0.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>0.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Ontario Ministry of Municipal Affairs and Housing, Financial Information Returns

21 In other parts of the country, social services and social housing are funded by the provincial level of government.
Manila

The Metropolitan Manila Development Authority lost some of its revenue sources with the passage of the Local Government Code in 1991. Under the Local Government Code, local government units were given the authority to impose a range of taxes including a property tax, tax on business operations, occupation tax, tax on real property transfers, tax on signs, signboards, billboards and other advertisements, franchise tax, amusement tax, cultural development tax, and residence tax. They were also given authority to collect service fees, user charges, and fines (Laquian 2002, 90).

As a result of these changes in the taxes that local government units could levy, the MMDA lost its 15 percent statutory share of the local government units’ regular income. Moreover, the funds from the constituent municipalities are often delayed or withheld. The MMDA revenues no longer include a share of real property taxes and its share of other local taxes has been reduced. As a result, the MMDA has become more dependent on central government grants. As Laquian (2002, 104) notes, the financial capabilities of the MMDA need to be strengthened if it is going to carry out metropolitan-wide functions.

Stuttgart

The Stuttgart Regional Association has no taxing authority and does not levy user fees. It is funded by municipal contributions (54 percent) and intergovernmental conditional grants from the Land of Baden-Württemburg (46 percent). The municipal funds consist of a general contribution (11 percent) and a contribution earmarked for public transit (35 percent). Both contributions are negotiated annually and split between the municipalities on the basis of tax raising capacity and structural factors (OECD 2006, 167). Revenues are also derived from transfers from county and municipal authorities (city-region administration levies plus a transport charge) and state transfers. The main expenditures of the association are regional express trains and the regional transport body that manages buses and tramways.
4. Concluding Comments

The population of urban areas around the world is growing, and it is growing more rapidly in less developed countries than in developed countries. With urban growth comes the challenge of delivering services, finding the resources to pay for those services, and coordinating service delivery across municipal boundaries. As the physical expansion of large cities extends into the rural hinterland and crosses into other political jurisdictions, there is an increasing need to coordinate service delivery. These challenges are particularly acute in developing countries where the population is poor and resources are limited.

How metropolitan areas are governed affects their ability to deliver and pay for services and to coordinate service delivery across municipal boundaries. Fragmentation reduces the ability of municipalities to solve regional problems on a regional basis and results in uneven fiscal resources and services. Coordination is especially important for those services that cross regional boundaries such as transportation, planning, and water and sewage systems. An effective system of governance for the entire metropolitan region is needed to ensure that services are delivered efficiently. Yet, few cities in the developing world (or the developed world) have an administrative structure at the metropolitan level.

This paper has reviewed metropolitan structures in many cities around the world. The findings of this review suggest that neither theory nor practice tells us clearly which model of governance is best for large metropolitan areas. The appropriate model depends on both the national and local context (constitution, division of responsibilities between central, state and local governments, sources of revenue at the local level, etc.). Models that work well in one jurisdiction may or may not be appropriate in another jurisdiction. Even in one jurisdiction, the model can and will change over time to reflect population growth, changes in the economy, or changes in the political situation. There is no “one size fits all” model of metropolitan governance.

In addition to the major finding that different models of governance work in different metropolitan areas around the world, the review of case studies in this paper highlights the following points:
• First, it is important to understand the national and local context of the metropolitan area, whether it means increased powers and responsibilities granted to municipalities in recent years in countries such as the Philippines and Brazil or the constitutional reforms in countries such as South Africa, Brazil, and India. This context will have a bearing on what is achievable. Where there is a strong tradition of local autonomy, such as in the Philippines for example, it will be more difficult to implement a metropolitan-wide structure that requires cooperation from local governments.

• Second, the governance structure should cover the entire economic region but this is rarely the case. Over time, economic boundaries evolve and change; political boundaries are much more difficult to modify. For this reason, there will likely be the need for some inter-municipal cooperation across municipal boundaries.

• Third, voluntary cooperation through inter-municipal agreements or special districts may be effective for the provision of some services but these approaches do not provide the needed regional foundation for metropolitan areas to coordinate service delivery. In situations in which local autonomy is paramount and where objectives are shared by policy-makers in various local governments, however, voluntary cooperation can work to some extent. It will not work well when objectives differ among local governments, as of course they often do.

• Fourth, the process of reforming the government structure is often as important as the outcome itself. All of the relevant stakeholders need to be included in the early stages of restructuring reform for the reform to be accepted. It is also important to understand that politics, and not economics, often drives the process and the outcome.

• Fifth, metropolitan governments need sufficient resources and fiscal autonomy to raise the funds they need to deliver services. If they do not have sufficient funds to carry out their mandate, they will be doomed to failure.
In conclusion, few problems and processes stop at municipal boundaries. The most feasible solutions require larger geographical units and access to a large pool of resources, both human and financial, than is likely to be at the disposal of small, local governments. A strong regional structure that encompasses the entire economic region is essential to ensure that services are delivered in a coordinated fashion across municipal boundaries.
References


